
MPSEARCH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Oct 4 15:24:16 1999; MasPar time 14.68 Seconds
44.916 Million cell updates/sec
ular output not generated.

Title: >MOHAM-312-CLAIM82A.PEP
Description: (1-31) from moham312177.pep
Perfect Score: 227
Sequence: 1 hgegtftsldskmqeeavrlfiewlknngp 31
Scoring table: PAM 150
Gap 11

Claim 82
X = P

Searched: 170751 segs, 21266608 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: a-geneseq35

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Statistics: Mean 24.388; Variance 104.481; scale 0.233

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	227	100.0	31	14	Heloderma suspectum e	3.13e-12
2	227	100.0	39	39	Exendin-4, for use in	3.13e-12
3	227	100.0	39	30	Gila monster exendin-	3.13e-12
4	227	100.0	39	14	Heloderma suspectum e	3.13e-12
5	227	100.0	87	35	Heloderma suspectum e	3.13e-12
6	220	96.9	39	39	Exendin-3, for use in	1.39e-11
7	220	96.9	39	14	Heloderma horridum ex	1.39e-11
8	220	96.9	39	30	Gila monster exendin-	1.39e-11
9	219	96.5	30	39	Exendin-4 (1-30) for	1.72e-11
10	219	96.5	31	14	Heloderma suspectum e	1.72e-11
11	212	93.4	30	29	H. horridum exendin-4	7.58e-11
12	207	91.2	30	29	H. horridum exendin-4	2.18e-10
13	205	90.3	28	39	Exendin-4 (1-28) amid	3.33e-10
14	205	90.3	30	29	H. horridum exendin-3	3.33e-10
15	205	90.3	30	29	H. horridum exendin-3	3.33e-10
16	203	89.4	30	29	H. horridum exendin-4	5.08e-10

17	202	89.0	39	39	W61773	Leu(14), Phe(25)-exen	6.27e-10
18	201	88.5	30	29	W39304	H. horridum exendin-4	7.74e-10
19	201	88.5	30	29	W39308	H. horridum exendin-4	7.74e-10
20	201	88.5	30	29	W39303	H. horridum exendin-4	7.74e-10
21	201	88.5	30	29	W39306	H. horridum exendin-4	7.74e-10
22	200	88.1	30	29	W39383	H. horridum exendin-3	9.55e-10
23	199	87.7	30	29	W39367	H. horridum exendin-4	1.18e-09
24	199	87.7	30	29	W39311	H. horridum exendin-4	1.18e-09
25	198	87.2	30	29	W39349	H. horridum exendin-4	1.45e-09
26	197	86.8	30	29	W39351	H. horridum exendin-4	1.80e-09
27	197	86.8	30	29	W39347	H. horridum exendin-4	1.80e-09
28	197	86.8	30	29	W39361	H. horridum exendin-4	1.80e-09
29	197	86.8	30	29	W39358	H. horridum exendin-4	1.80e-09
30	196	86.3	30	29	W39341	H. horridum exendin-4	2.21e-09
31	196	86.3	30	29	W39345	H. horridum exendin-4	2.21e-09
32	196	86.3	30	29	W39310	H. horridum exendin-4	2.21e-09
33	195	85.9	30	29	W39317	H. horridum exendin-4	2.73e-09
34	195	85.9	30	29	W39331	H. horridum exendin-4	2.73e-09
35	195	85.9	30	29	W39343	H. horridum exendin-3	3.37e-09
36	194	85.5	30	29	W39305	H. horridum exendin-3	3.37e-09
37	194	85.5	30	29	W39319	H. horridum exendin-3	3.37e-09
38	194	85.5	30	29	W39420	H. horridum exendin-3	3.37e-09
39	194	85.5	30	29	W39369	H. horridum exendin-3	3.37e-09
40	194	85.5	30	29	W39370	H. horridum exendin-3	3.37e-09
41	194	85.5	30	29	W39378	H. horridum exendin-3	3.37e-09
42	194	85.5	30	29	W39327	H. horridum exendin-4	3.37e-09
43	193	85.0	30	29	W39332	H. horridum exendin-4	4.16e-09
44	193	85.0	30	29	W39354	H. horridum exendin-4	4.16e-09
45	193	85.0	30	29	W39380	H. horridum exendin-3	4.16e-09

ALIGNMENTS

RESULT 1
ID R80543 standard; peptide; 31 AA.

AC R80543; 27-FEB-1996 (first entry)

DE Heloderma suspectum exendin-4 residues 1-31 (Exendin-4(1-31)).
KW Exendin-4; residues 1-31; Exendin-4(1-31); diabetes mellitus;
hyperglycaemia; insulinotropic peptide.

OS Heloderma suspectum.

PN US5424286-A.

FD 13-JUN-1995.

PR 24-MAY-1993; 066480.

PF 24-MAY-1993; US-066480.

PA (ENGJ/) ENG J.

PI Eng J.

DR WPI; 95-262627/34.

PT Stimulating/inhibiting insulin release with exendin polypeptide(s) -
for treating diabetes mellitus and preventing hyperglycaemia.

PS Claim 1; Columns 13-14; 17pp; English.

CC R80543 is the Heloderma suspectum exendin-4 residues 1-31. It is an
insulinotropic peptide, and can therefore be used in the treatment of
diabetes mellitus (types I or II), and for the prevention of
hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
CC and insulin-(in)dependent mechanisms.

SC Sequence 31 AA;

Query Match 100.0%; Score 227; DB 14; Length 31;

Best Local Similarity 100.0%; Pred. No. 3.13e-12;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtftsldskmqeeavrlfiewlknngp 31

|||||

Qy 1 hgegtftsldskmqeeavrlfiewlknngp 31

RESULT 2

ID W61770 standard; peptide; 39 AA.

AC W61770;

DT 29-MAR-1999 (first entry)

DE Exendin-4, for use in treating disorders related to food intake.

KW Exendin; obesity; type II diabetes; eating disorders; cardiac disease;

KW insulin resistance syndrome; elevated plasma glucose level; agonist.
 OS Heloderma suspectum.
 PN WO9830231-A1.
 PD 16-JUL-1998.
 PF 07-JAN-1998; U00449.
 PR 14-NOV-1997; US-066029.
 PR 07-JAN-1997; US-034905.
 PR 08-AUG-1997; US-055404.
 PR 14-NOV-1997; US-065442.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Beeley NRA, Bhavsar S, Prickett KS;
 DR WPI; 98-398796/34.
 PT Reducing food intake by administering extendins or their
 PT analogues - for treatment of e.g. obesity, type II diabetes,
 PT eating disorders and insulin resistance
 PS Claims 17, 25; Page 8; 21pp; English.
 The invention relates to a new method for treating disorders that
 are alleviated by reducing food intake, in particular obesity, type
 II diabetes, eating disorders, insulin resistance syndrome, elevated
 plasma glucose levels, or the risk of cardiac disease. The method
 comprises administering an extendin or an extendin agonist. The treatment
 reduces appetite and lowers plasma lipid levels. It inhibits food
 consumption as effectively as amylin or cholecystokinin but has a much
 longer-lasting action (still effective after 6 hours in a mouse model).
 The present sequence is that of extendin-4 which is one of the preferred
 compounds for use in the method.
 SQ Sequence 39 AA;

Query Match 100.0%; Score 227; DB 39; Length 39;
 Best Local Similarity 100.0%; Pred. No. 3.13e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgeftfslskmqmeeeavrflfiewlknngp 31
 |||||
 Qy 1 hgeftfslskmqmeeeavrflfiewlknngp 31

RESULT 3
 ID W47609 standard; peptide; 39 AA.
 AC W47609;
 DT 03-JUL-1998 (first entry)
 DE Gila monster extendin-4.
 KW Extendin agonist; gastric motility; gastric emptying; treatment;
 PT spasms; postprandial dumping syndrome; postprandial hyperglycaemia;
 PT type 1 diabetes; impaired glucose tolerance; toxin ingestion;
 KW obesity; Gila monster venom; extendin-4.
 OS Heloderma suspectum.
 Key Location/Qualifiers
 Modified_site 39 /note= "amidated"

WO9805351-A1.
 12-FEB-1998.
 PF 08-AUG-1997; U14199.
 PR 08-AUG-1996; US-694954.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Beeley NRA, Gedulin B, Prickett KS, Young AA;
 DR WPI; 98-145351/13.
 PT Regulating gastrointestinal motility using extendins or their
 PT agonists - for treating spasm, diabetic postprandial hyperglycaemia,
 PT impaired glucose tolerance etc. also in diagnostic investigations
 PS Claims 20 and 21; Fig 1; 70pp; English.
 CC W47549 describes a generic extendin agonist, provided that it does
 CC have the formula of either extendin-3 (W47608) or extendin-4
 CC (W47609).
 CC Extendin agonists, which reduce gastric motility and delay gastric
 CC emptying, can be used to treat spasm (where associated with acute
 CC diverticulitis or disorders of the biliary tract or sphincter of
 CC Oddi), postprandial dumping syndrome and hyperglycaemia
 CC (particularly associated with type 2 diabetes), type 1 diabetes,
 CC impaired glucose tolerance, toxin ingestion (an extendin agonist is
 CC administered to prevent stomach contents passing into the
 CC intestines, then the stomach pumped) and obesity. They can also be
 CC administered to subjects undergoing gastrointestinal diagnostic

CC investigation, particularly radiological or by magnetic resonance
 CC imaging.
 CC Extendins, components of Gila monster venom, have some sequence
 CC similarity to glucagon-like peptides (GLP). They are GLP agonists
 CC and have been suggested (US5424286) for treatment of diabetes and
 CC prevention of hyperglycaemia.
 SQ Sequence 39 AA;

Query Match 100.0%; Score 227; DB 30; Length 39;
 Best Local Similarity 100.0%; Pred. No. 3.13e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgeftfslskmqmeeeavrflfiewlknngp 31
 |||||
 Qy 1 hgeftfslskmqmeeeavrflfiewlknngp 31

RESULT 4
 ID R80546 standard; peptide; 39 AA.
 AC R80546;
 DT 27-FEB-1996 (first entry)
 DE Heloderma suspectum extendin-4.
 KW Extendin-4; diabetes mellitus; hyperglycaemia; insulinotropic peptide.
 OS Heloderma suspectum.
 PN US5424286-A.
 PD 13-JUN-1995.
 PF 24-MAY-1993; 066480.
 PR 24-MAY-1993; US-066480.
 PA (ENGJ/) ENG J.
 PI Eng J;
 DR WPI; 95-262627/34.
 PT Stimulating/inhibiting insulin release with extendin polypeptide(s) -
 PT for treating diabetes mellitus and preventing hyperglycaemia.
 PS Claim 6; Columns 13-14; 17pp; English.
 CC R80546 is Heloderma suspectum extendin-4. It is an
 CC insulinotropic peptide, and can therefore be used in the treatment of
 CC diabetes mellitus (types I or II), and for the prevention of
 CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
 CC and insulin-(in)dependent mechanisms.
 SQ Sequence 39 AA;

Query Match 100.0%; Score 227; DB 14; Length 39;
 Best Local Similarity 100.0%; Pred. No. 3.13e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgeftfslskmqmeeeavrflfiewlknngp 31
 |||||
 Qy 1 hgeftfslskmqmeeeavrflfiewlknngp 31

RESULT 5
 ID W70288 standard; Protein; 87 AA.
 AC W70288;
 DT 06-NOV-1998 (first entry)
 DE Heloderma suspectum proextendin peptide.
 KW Heloderma suspectum proextendin; extendin N-terminal peptide; ENTP;
 KW extendin 4 peptide; extendin 3 gene; Heloderma horridum; metabolic disease;
 KW drug screening; endocrine tumour; organ failure; cell metabolism;
 KW diabetes; reptilian venom peptide.
 OS Heloderma suspectum.
 FH Key Location/Qualifiers
 FT Peptide 1..23 /note= "Signal peptide"
 FT Peptide 1..47 /note= "ENTP"
 FT Peptide 48..87 /note= "Extendin 4"
 FT Cleavage_site 46..47 /note= "Dipeptidyl peptidase cleavage site"
 PN WO9835033-A1.
 PD 13-AUG-1998.
 PF 04-FEB-1998; CA0071.
 PR 07-FEB-1997; GB-002582.

PR 05-FEB-1997; US-037412.
 PA (OMEO-) 1149336 ONTARIO INC.
 DR Drucker DJ;
 WI WPI: 98-447230/38.
 N-2SDB: V33163.

PT New nucleic acid encoding proexendin - used to diagnose and treat,
 e.g. endocrine tumours, also to treat poisoning by reptile venom
 Claim 3; Fig 2; 26pp; English.
 CC The Heloderma suspectum proexendin peptide is encoded by its cDNA
 CC which was isolated from a H. suspectum salivary gland cDNA library.
 CC The proexendin protein comprises of a novel exendin N-terminal
 CC peptide (PNTP) linked to the N-terminus of the exendin 4 peptide
 CC by a consensus dipeptidyl peptidase cleavage site. The proexendin
 CC cDNA can be used to clone or identify related sequences (e.g. the
 CC exendin 3 gene of Heloderma horridum, mutant alleles and proexendin
 CC gene regulatory defects associated with metabolic disease) and species
 CC homologues (e.g. for developing animal models for drug screening).
 CC The proexendin peptide can be used to raise antibodies. Anti-proexendin
 CC antibodies are claimed to be useful for diagnosing conditions associated
 CC with altered levels of proexendin (e.g. endocrine tumours and organ
 CC failure), for identifying other regulators of cell metabolism, in drug
 CC screens and for treating metabolic diseases (e.g. diabetes) and for
 CC neutralising, or detecting, reptilian venom peptides.
 CC Sequence 87 AA;

Query Match 100.0%; Score 227; DB 35; Length 87;
 Best Local Similarity 100.0%; Pred. No. 3 13e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 48 hgegtftsldskmqeeavrlfiewlknngpp 78

QY 1 hgegtftsldskmqeeavrlfiewlknngpp 31

RESULT 6

ID W61769 standard; peptide; 39 AA.

AC W61769;
 DT 29-MAR-1999 (first entry)
 DE Exendin-3, for use in treating disorders related to food intake.
 KW Exendin; obesity; type II diabetes; eating disorders; cardiac disease;
 KW insulin resistance syndrome; elevated plasma glucose level; agonist.
 OS Heloderma horridum.
 PS WO98030231-A1.
 PN 16-JUL-1998.
 PF 07-JAN-1998; U00449.
 PR 14-NOV-1997; US-066029.
 PR 07-JAN-1997; US-034905.
 PR 08-AUG-1997; US-055404.
 PR 14-NOV-1997; US-065442.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Beeley NRA, Bhavsar S, Prickett KS;
 WPI: 98-398796/34.
 PT Reducing food intake by administering exendins or their
 PT analogues - for treatment of e.g. obesity, type II diabetes,
 PT eating disorders and insulin resistance
 PS Claims 16, 24; Page 8; 214pp; English.
 CC The invention relates to a new method for treating disorders that
 CC are alleviated by reducing food intake, in particular obesity, type
 CC II diabetes, eating disorders, insulin resistance syndrome, elevated
 CC plasma glucose levels, or the risk of cardiac disease. The method
 CC comprises administering an exendin or an exendin agonist. The treatment
 CC reduces appetite and lowers plasma lipid levels. It inhibits food
 CC consumption as effectively as amylin or cholecystokinin but has a much
 CC longer-lasting action (still effective after 6 hours in a mouse model).
 CC The present sequence is that of exendin-3 which is one of the preferred
 CC compounds for use in the method.
 CC Sequence 39 AA;

Query Match 96.9%; Score 220; DB 39; Length 39;
 Best Local Similarity 93.5%; Pred. No. 1 39e-11;
 Matches 29; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 hsdgtftsldskmqeeavrlfiewlknngpp 31

QY 1 hgegtftsldskmqeeavrlfiewlknngpp 31

RESULT 7

ID R80545 standard; peptide; 39 AA.

AC R80545;
 DT 27-FEB-1996 (first entry)
 DE Heloderma horridum exendin-3.
 KW Exendin-3; diabetes mellitus; hyperglycaemia; insulinotropic peptide.
 OS Heloderma horridum.
 PN US5424286-A.
 PF 13-JUN-1993.
 PF 24-MAY-1993; 066480.
 PR 24-MAY-1993; US-066480.
 PA (ENGJ/) ENG J.
 PI Eng J;
 WI WPI: 95-262627/34.
 DT Stimulating/inhibiting insulin release with exendin polypeptide(s) -
 DT for treating diabetes mellitus and preventing hyperglycaemia.
 PS Claim 5; Columns 13-14; 17pp; English.
 CC R80545 is Heloderma horridum exendin-3. It is an
 CC insulinotropic peptide, and can therefore be used in the treatment of
 CC diabetes mellitus (types I or II), and for the prevention of
 CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
 CC and insulin-(in)dependent mechanisms.
 CC Sequence 39 AA;

Query Match 96.9%; Score 220; DB 14; Length 39;
 Best Local Similarity 93.5%; Pred. No. 1 39e-11;
 Matches 29; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 hsdgtftsldskmqeeavrlfiewlknngpp 31

QY 1 hgegtftsldskmqeeavrlfiewlknngpp 31

RESULT 8

ID W47608 standard; peptide; 39 AA.

AC W47608;
 DT 03-JUL-1998 (first entry)
 DE Gila monster exendin-3.
 KW Exendin agonist; gastric motility; gastric emptying; treatment;
 KW spasm; postprandial dumping syndrome; postprandial hyperglycaemia;
 KW type I diabetes; impaired glucose tolerance; toxin ingestion;
 KW obesity; Gila monster venom; exendin-3.
 OS Heloderma horridum.

Key Location/Qualifiers

FT Modified_site 39 /note= "amidated"

FT WO9805351-A1.

PN 12-FEB-1998.

PF 08-AUG-1997; U14199.

PR 08-AUG-1996; US-694954.

PA (AMYL-) AMYLIN PHARM INC.

PI Beeley NRA, Gedulin B, Prickett KS, Young AA;

WPI: 98-145351/13.

DR Regulating gastrointestinal motility using exendins or their
 DR agonists - for treating spasm, diabetic postprandial hyperglycaemia,
 PT impaired glucose tolerance etc., also in diagnostic investigations
 PT Claims 20 and 21; Fig 1, 70pp; English.
 PS W47349 describes a generic exendin agonist, provided that it does
 CC have the formula of either exendin-3 (W47608) or exendin-4
 CC (W47609).

CC Exendin agonists, which reduce gastric motility and delay gastric
 CC emptying, can be used to treat spasm (where associated with acute
 CC diverticulitis or disorders of the biliary tract or sphincter of
 CC Oddi), postprandial dumping syndrome and hyperglycaemia
 CC (particularly associated with type 2 diabetes), type 1 diabetes,
 CC impaired glucose tolerance, toxin ingestion (an exendin agonist is
 CC administered to prevent stomach contents passing into the
 CC intestines, then the stomach pumped) and obesity. They can also be
 CC administered to subjects undergoing gastrointestinal diagnostic

CC investigation, particularly radiological or by magnetic resonance
 CC imaging.
 CC Extensins, components of Gila monster venom, have some sequence
 CC similarity to glucagon-like peptides (GLP). They are GLP agonists
 CC and have been suggested (US5424286) for treatment of diabetes and
 CC prevention of hyperglycaemia.
 CC Sequence 39 AA;

Query Match 96.9%; Score 220; DB 30; Length 39;
 Best Local Similarity 93.5%; Pred. No. 1.39e-11;
 Matches 29; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 hsgdftsdlskqmeeeavrlfiewlknngp 31
 |:|||||:|||||:|||||:|||||:|||||
 QY 1 hgegtftsdlkqmeeeavrlfiewlknngp 31

ULT 9
 W61771 standard; peptide; 30 AA.

DT 29-MAR-1999 (first entry)
 DE Exendin-4 (1-30) for use in treating disorders related to food intake.
 KW Exendin; obesity; type II diabetes; eating disorders; cardiac disease;
 KW insulin resistance syndrome; elevated plasma glucose level; agonist.
 CS Heloderma suspectum.

Key Location/Qualifiers
 Modified_site 30

PN W09830231-A1. /note= "Optionally the C-terminal is in amide form"

PD 16-JUL-1998.

PF 07-JAN-1998; 000449.

PR 14-NOV-1997; US-066029.

PR 07-JAN-1997; US-034905.

PR 08-AUG-1997; US-055404.

PR 14-NOV-1997; US-065442.

PA (AMYL-) AMYLIN PHARM INC.

PI Beely NRA, Bhavsar S, Prickett KS;

PT WPI; 98-398796/34.

PT Reducing food intake by administering exendins or their
 analogues - for treatment of e.g. obesity, type II diabetes,
 eating disorders and insulin resistance

PS Claims 18, 26; Page 11; 214pp; English.

CC The invention relates to a new method for treating disorders that
 are alleviated by reducing food intake, in particular obesity, type
 II diabetes, eating disorders, insulin resistance syndrome, elevated
 plasma glucose levels, or the risk of cardiac disease. The method
 comprises administering an exendin or an exendin agonist. The treatment
 reduces appetite and lowers plasma lipid levels. It inhibits food
 consumption as effectively as amylin or cholecystokinin but has a much
 longer-lasting action (still effective after 6 hours in a mouse model).

CC The present sequence is that of exendin-4 (1-30) or its amide which is
 CC one of the preferred compounds for use in the method.
 CC Sequence 30 AA;

Query Match 96.5%; Score 219; DB 39; Length 30;
 Best Local Similarity 100.0%; Pred. No. 1.72e-11;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtftsdlkqmeeeavrlfiewlknng 30
 |:|||||:|||||:|||||:|||||:|||||
 QY 1 hgegtftsdlkqmeeeavrlfiewlknng 30

RESULT 10

ID R80544 standard; peptide; 31 AA.

AC R80544;

DT 27-FEB-1996 (first entry)

DE Heloderma suspectum exendin-4 residues 1-31-Tyr31.

KW Exendin-4; residues 1-31; Y-31-Exendin-4(1-31); diabetes mellitus;
 KW hyperglycaemia; Tyr31; insulinotropic peptide.

OS Heloderma suspectum.

PN US5424286-A.

PD 13-JUN-1995.

PF 24-MAY-1993; 066480.
 PR 24-MAY-1993; US-066480.
 PA (ENGJ/) ENG J.

PI Eng J;

DR WPI; 95-262627/34.

PT Stimulating/inhibiting insulin release with exendin polypeptide(s) -
 for treating diabetes mellitus and preventing hyperglycaemia.

PS Claim 2; Columns 13-14; 17pp; English.

CC R80544 is the Heloderma suspectum exendin-4 residues 1-31, where
 CC the native Pro31 has been replaced with a Tyr residue. It is an

CC insulinotropic peptide, and can therefore be used in the treatment of

CC diabetes mellitus (types I or II), and for the prevention of

CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent

CC and insulin-(in)dependent mechanisms.

CC Sequence 31 AA;

Query Match 96.5%; Score 219; DB 14; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.72e-11;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtftsdlkqmeeeavrlfiewlknng 30
 |:|||||:|||||:|||||:|||||:|||||

QY 1 hgegtftsdlkqmeeeavrlfiewlknng 30

RESULT 11

ID W39302 standard; peptide; 30 AA.

AC W39302;

DT 05-JUN-1998 (first entry)

DE H. horridum exendin-4 peptide.

KW Exendin-3; exendin 4; diabetes; insulin; secretion; biosynthesis;

KW glucagon reduction; hypoglycaemia; glucose; treatment.

OS Heloderma horridum.

Key Location/Qualifiers

Modified_site 30

FT /note= "This residue can be any amino acid except
 FT Gly"

PN W09746584-A1.

PD 11-DEC-1997.

PF 05-JUN-1997; R02930.

PR 13-SEP-1996; DE-037230.

PR 05-JUN-1996; DE-022502.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Goetze B, Goetze R, Hoffmann E;

DR WPI; 98-042119/04.

PT Truncated versions of exendin peptide(s) for treating diabetes -

PT increase secretion and biosynthesis of insulin, but reduce those of

PT glucagon, and do not induce hypoglycaemia

PS Claim 1; Page 4; 150pp; English.

CC This peptide is a fragment of exendin-4 isolated from Heloderma

CC horridum. This peptide and its salts, esters and derivatives can be

CC used to treat diabetes mellitus. They stimulate biosynthesis and

CC secretion of insulin, but have the opposite effect on glucagon, and

CC independent of this activity can increase peripheral glucose utilisation.

CC Exendin-3 and exendin-4 are only active when blood sugar levels are

CC high, so they will not induce hypoglycaemia. Compared with glucagon-like

CC peptide 1 (GLP1) and the known exendins, they are more active (effective

CC at lower doses), more stable to degradation and metabolism and have a

CC longer lasting effect. Truncated forms of this peptide can be made more

CC economically than full length versions.

CC Sequence 30 AA;

Query Match

Best Local Similarity 93.4%; Score 212; DB 29; Length 30;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtftsdlkqmeeeavrlfiewlknng 29
 |:|||||:|||||:|||||:|||||:|||||

QY 1 hgegtftsdlkqmeeeavrlfiewlknng 29

RESULT 12

ID W39309 standard; peptide; 30 AA.

AC W39309;
 DT 05-JUN-1998 (first entry)
 DE H. horridum extendin-4 peptide derivative #6.
 KW Extendin-3; extendin 4; diabetes; insulin; secretion; biosynthesis;
 KW glucagon reduction; hypoglycaemia; glucose; treatment.
 OS Heloderma horridum.
 FH Key Location/Qualifiers
 FT Modified_site 30 /note= "C-terminal amide"
 FT W09746584-A1.
 PN 11-DEC-1997.
 PF 05-JUN-1997; E02930.
 PR 13-SEP-1996; DE-037230.
 PR 05-JUN-1996; DE-022502.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Goeke B, Goeke R, Hoffmann E;
 DR WPI; 98-042119/04.
 DR Truncated versions of extendin peptide(s) for treating diabetes -
 PT increase secretion and biosynthesis of insulin, but reduce those of
 PT glucagon, and do not induce hypoglycaemia
 PS Claim 2; Page 22; 15Opp; English.
 CC Peptides W39303-W39420 are fragments of extendin-3 and extendin-4
 CC isolated from Heloderma horridum which are used in a novel method
 CC for the treatment of diabetes mellitus. These peptides can stimulate
 CC biosynthesis and secretion of insulin, but have the opposite effect on
 CC glucagon, and independent of this activity can increase peripheral
 CC glucose utilisation. Extendin-3 and extendin-4 are only active when blood
 CC sugar levels are high, so they will not induce hypoglycaemia. Compared
 CC with glucagon-like peptide 1 (GLP1) and the known extendins, they are more
 CC active (effective at lower doses), more stable to degradation and
 CC metabolism and have a longer lasting effect. Truncated forms of this
 CC peptide can be made more economically than full length versions.
 SQ Sequence 30 AA;
 Query Match 91.28; Score 207; DB 29; Length 30;
 Best Local Similarity 96.63; Pred. No. 2.18e-10;
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 1 hgegttsdlskqmeeeavrlfiewlkn 29
 QY 1 hgegttsdlskqmeeeavrlfiewlkn 29
 RESULT 13
 ID W61772 standard; peptide; 28 AA.
 AC W61772;
 DT 29-MAR-1999 (first entry)
 DE Extendin-4 (1-28) amide for use in treating disorders related to food.
 KW Extendin; obesity; type II diabetes; eating disorders; cardiac disease;
 KW insulin resistance syndrome; elevated plasma glucose level; agonist.
 OS Heloderma suspectum.
 FH Key Location/Qualifiers
 FT Modified_site 28 /note= "the C-terminal is in amide form"
 FT W09830231-A1.
 PN 16-JUL-1998.
 PF 07-JAN-1998; U00449.
 PR 14-NOV-1997; US-066029.
 PR 07-JAN-1997; US-034905.
 PR 08-AUG-1997; US-055404.
 PR 14-NOV-1997; US-065442.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Beekley NRA, Bhavsar S, Prickett KS;
 DR WPI; 98-398796/34.
 DR Reducing food intake by administering extendins or their
 PT analogues - for treatment of e.g. obesity, type II diabetes,
 PT eating disorders and insulin resistance
 PS Claims 18, 26; Page 12; 214pp; English.
 CC The invention relates to a new method for treating disorders that
 CC are alleviated by reducing food intake, in particular obesity, type
 CC II diabetes, eating disorders, insulin resistance syndrome, elevated
 CC plasma glucose levels, or the risk of cardiac disease. The method
 CC comprises administering an extendin or an extendin agonist. The treatment

CC reduces appetite and lowers plasma lipid levels. It inhibits food
 CC consumption as effectively as amylin or cholecystokinin but has a much
 CC longer-lasting action (still effective after 6 hours in a mouse model).
 CC The present sequence is that of extendin-4 (1-28) amide which is one of
 CC the preferred compounds for use in the method.
 SQ Sequence 28 AA;
 Query Match 90.3%; Score 205; DB 39; Length 28;
 Best Local Similarity 100.0%; Pred. No. 3.33e-10;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 hgegttsdlskqmeeeavrlfiewlkn 28
 QY 1 hgegttsdlskqmeeeavrlfiewlkn 28
 RESULT 14
 ID W39368 standard; peptide; 30 AA.
 AC W39368;
 DT 05-JUN-1998 (first entry)
 DE H. horridum extendin-3 peptide derivative #11.
 KW Extendin-3; extendin 4; diabetes; insulin; secretion; biosynthesis;
 KW glucagon reduction; hypoglycaemia; glucose; treatment.
 OS Heloderma horridum.
 FH Key Location/Qualifiers
 FT Modified_site 30 /note= "C-terminal amide"
 FT W09746584-A1.
 PN 11-DEC-1997.
 PF 05-JUN-1997; E02930.
 PR 13-SEP-1996; DE-037230.
 PR 05-JUN-1996; DE-022502.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Goeke B, Goeke R, Hoffmann E;
 DR WPI; 98-042119/04.
 DR Truncated versions of extendin peptide(s) for treating diabetes -
 PT increase secretion and biosynthesis of insulin, but reduce those of
 PT glucagon, and do not induce hypoglycaemia
 PS Claim 2; Page 27; 15Opp; English.
 CC Peptides W39303-W39420 are fragments of extendin-3 and extendin-4
 CC isolated from Heloderma horridum which are used in a novel method
 CC for the treatment of diabetes mellitus. These peptides can stimulate
 CC biosynthesis and secretion of insulin, but have the opposite effect on
 CC glucagon, and independent of this activity can increase peripheral
 CC glucose utilisation. Extendin-3 and extendin-4 are only active when blood
 CC sugar levels are high, so they will not induce hypoglycaemia. Compared
 CC with glucagon-like peptide 1 (GLP1) and the known extendins, they are more
 CC active (effective at lower doses), more stable to degradation and
 CC metabolism and have a longer lasting effect. Truncated forms of this
 CC peptide can be made more economically than full length versions.
 SQ Sequence 30 AA;
 Query Match 90.3%; Score 205; DB 29; Length 30;
 Best Local Similarity 93.1%; Pred. No. 3.33e-10;
 Matches 27; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Db 1 hsdgttsdlskqmeeeavrlfiewlkn 29
 QY 1 hgegttsdlskqmeeeavrlfiewlkn 29
 RESULT 15
 ID W39301 standard; peptide; 30 AA.
 AC W39301;
 DT 05-JUN-1998 (first entry)
 DE H. horridum extendin-3 peptide.
 KW Extendin-3; extendin 4; diabetes; insulin; secretion; biosynthesis;
 KW glucagon reduction; hypoglycaemia; glucose; treatment.
 OS Heloderma horridum.
 FH Key Location/Qualifiers
 FT Modified_site 30 /note= "This residue can be any amino acid except
 FT Gly"

PN WO9746584-A1.
 PD 11-DEC-1997.
 PF 05-JUN-1997; E02930.
 PR 13-SEP-1996; DE-037230.
 PK 05-JUN-1996; DE-022502.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Goeke B, Goeke R, Hoffmann E;
 DR WPI; 98-042119/04.
 PT Truncated versions of exendin peptide(s) for treating diabetes -
 PT increase secretion and biosynthesis of insulin, but reduce those of
 PT glucagon, and do not induce hypoglycaemia
 PS Claim 1, Page 3; 150pp; English.
 CC This peptide is a fragment of exendin-3 isolated from Heloderma
 CC horridum. This peptide and its salts, esters and derivatives can be
 CC used to treat diabetes mellitus. They stimulate biosynthesis and
 CC secretion of insulin, but have the opposite effect on glucagon, and
 CC independent of this activity can increase peripheral glucose utilisation.
 CC Exendin-3 and exendin-4 are only active when blood sugar levels are
 CC high, so they will not induce hypoglycaemia. Compared with glucagon-like
 CC peptide 1 (GLP1) and the known exendins, they are more active (effective
 CC at lower doses), more stable to degradation and metabolism and have a
 CC longer lasting effect. Truncated forms of this peptide can be made more
 CC economically than full length versions.
 SQ Sequence 30 AA;

Query Match 90.3%; Score 205; DB 29; Length 30;
 Best Local Similarity 93.1%; Pred. No. 3.33e-10;
 Matches 27; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 hsdgtfstdlskameeeavrffiewlknk 29
 QY 1 hgegtfstdlskameeeavrffiewlknk 29

Search completed: Mon Oct 4 15:24:34 1999
 Job time : 18 secs.

W P S R E H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Oct 4 15:23:50 1999; MasPar time 6.03 Seconds
206.071 Million cell updates/sec
ular output not generated.

--file: >MOHAM-312-CLAIM82A.PEP
Description: (1-31) from moham312177.pep
Perfect Score: 227
Sequence: 1 hgegtftsldskmqeeavrlfiewlknngpp 31

Scoring table: PAM 150
Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: p160
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 34.029; Variance 63.644; scale 0.535

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	227	100.0	39	1	HWGH4G	
2	220	96.9	39	1	HWGH3Z	exendin-4 - Gila mons 3.06e-27
3	127	55.9	31	2	S44472	exendin-3 - Mexican b 7.58e-26
4	127	55.9	101	1	GCFGB	glucagon G2 - North A 2.68e-08
5	125	55.1	63	1	GCIDC	glucagon precursor - 2.68e-08
6	122	53.7	30	2	S44473	glucagon-like peptide 5.97e-08
7	122	53.7	31	2	S44471	glucagon G1 - North A 1.97e-07
8	121	53.3	30	2	B61125	glucagon-like peptide 2.93e-07
9	121	53.3	30	2	C61125	glucagon-like peptide 2.93e-07
10	120	52.9	66	2	I51093	glucagon - chnook sa 4.35e-07
11	120	52.9	178	2	I51057	glucagon I precursor 4.35e-07
12	120	52.9	178	2	I51058	glucagon I precursor 4.35e-07
13	119	52.4	72	1	GCGBA	glucagon precursor - 6.45e-07
14	118	52.0	122	1	GCAF2	glucagon 2 precursor 9.55e-07
15	117	51.5	60	1	GCNC	glucagon precursor - 1.41e-06
16	115	50.7	29	1	GCDF	glucagon - smaller sp 3.08e-06
17	115	50.7	158	1	CGPG	glucagon precursor - 3.08e-06
18	115	50.7	180	2	A57294	glucagon precursor - 3.08e-06
19	115	50.7	180	1	GCHY	glucagon precursor - 3.08e-06
20	115	50.7	180	1	GCRTDU	glucagon precursor - 3.08e-06
21	115	50.7	180	1	GCBO	glucagon precursor - 3.08e-06
22	115	50.7	180	1	GCGP	glucagon precursor - 3.08e-06
23	115	50.7	180	1	GCHU	glucagon precursor - 3.08e-06

24 115 50.7 180 1 GCRT glucagon precursor - 3.08e-06
25 114 50.2 151 1 GCHU glucagon precursor - 4.54e-06
26 114 50.2 206 2 I51301 proglucagon - chicken 4.54e-06
27 113 49.8 124 1 GCRF glucagon 1 precursor 6.88e-06
28 110 48.5 29 1 GCBF glucagon - Chinchilla 2.11e-05
29 108 47.6 29 2 S07211 glucagon - marbled el 4.52e-05
30 104 45.8 29 1 GCFLE glucagon - European f 2.04e-04
31 104 45.8 29 2 A61135 glucagon - bigeye tun 2.04e-04
32 104 45.8 87 1 GCFIS glucagon precursor - 2.04e-04
33 100 44.1 29 2 A31742 glucagon - Arabian ca 8.96e-04
34 100 44.1 29 2 A31741 glucagon - rabbit (te 8.96e-04
35 100 44.1 29 2 C39258 glucagon - common squ 8.96e-04
36 100 44.1 69 1 GCDG69 glucagon-69 - dog 1.29e-03
37 99 43.6 29 2 S39018 glucagon - bowfin 1.29e-03
38 99 43.6 29 1 GCEN glucagon - elephantfi 1.86e-03
39 98 43.2 29 2 C50840 glucagon I - European 1.86e-03
40 96 42.3 29 1 GCOPV glucagon - North Amer 3.84e-03
41 96 42.3 29 2 A91740 glucagon - turkey (te 3.84e-03
42 95 41.9 29 1 A61583 glucagon - ostrich 5.50e-03
43 95 41.9 29 1 GCDTS glucagon - slider tur 5.50e-03
44 95 41.9 29 1 GCDK glucagon - duck 5.50e-03
45 91 40.1 36 1 GCFI glucagon-36 - spotted 2.28e-02

ALIGNMENTS

RESULT 1
ENTRY HWGH4G #type complete
TITLE exendin-4 - Gila monster
ORGANISM #formal_name Heloderma suspectum #common_name Gila monster
DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 21-Nov-1997
ACCESSIONS A42486
REFERENCE A42486
#authors Eng, J.; Kleinman, W.A.; Singh, L.; Singh, G.; Raufman, J.P.
#journal J. Biol. Chem. (1992) 267:7402-7405
#title Isolation and characterization of exendin-4, an exendin-3 analogue, from Heloderma suspectum venom. Further evidence for an exendin receptor on dispersed acini from guinea pig pancreas.
#cross-references MUID:92218391
#accession A42486
#molecule_type protein
#residues 1-39 #label ENG
COMMENT Exendin-4 does not stimulate amylase secretion by pancreatic acinar cells.

CLASSIFICATION #superfamily glucagon
KEYWORDS amidated carboxyl end; duplication; venom
FEATURE 39
#modified_site amidated carboxyl end (Ser) #status experimental
SUMMARY #length 39 #molecular-weight 4188 #checksum 9570
Query Match 100.0%; Score 227; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.06e-27;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 HGEFTTSLDKMQEEAVRLFIWLNKNGP 31
QY 1 hgegtftsldskmqeeavrlfiewlknngpp 31

RESULT 2
ENTRY HWGH3Z #type complete
TITLE exendin-3 - Mexican beaded lizard
ORGANISM #formal_name Heloderma horridum #common_name Mexican beaded lizard
DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 21-Nov-1997
ACCESSIONS A23674
REFERENCE A23674
#authors Eng, J.; Andrews, P.C.; Kleinman, W.A.; Singh, L.; Raufman, J.P.

```

J. Biol. Chem. (1985) 263:9746-9751
Isolation of peptide hormones from the pancreas of the
bullfrog (Rana catesbeiana). Amino acid sequences of
pancreatic polypeptide, oxyntomodulin, and two
glucagon-like peptides.
#cross-references MUID:80257102
#accession B28091
##molecule_type protein
##residues 1-36 ##label PO2
#accession C28091
##molecule_type protein
##residues 37-68 ##label POL
#accession D28091
##molecule_type protein
##residues 69-101 ##label PO3
CLASSIFICATION
KEYWORDS superfamily glucagon
FEATURE carbohydrate metabolism; duplication; hormone; pancreas
1-36      #product glucagon-36 (oxyntomodulin) #status
          experimental #label G36\
1-29      #product glucagon #status predicted #label GCN\
37-67     #product glucagon-like peptide 1 #status experimental
          #label GL1\
        69-101    #product glucagon-like peptide 2 #status experimental
          #label GL2
SUMMARY   #length 101 #checksum 9108

Query Match       55.9%; Score 127; DB 1; Length 101;
Best Local Similarity 51.6%; Pred. No. 2.68e-08;
Matches 16; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 37 HADGFTSDMSSYLEEAKAEFVDWLKGRP 67
|::|||::|::|::|::|::|::|::|::|::|
QY 1 hgegtfslskmqeeavrlfiewlknngp 31

RESULT 5
ENTRY GCDC #type fragments
TITLE glucagon precursor - channel catfish (fragments)
ORGANISM #formal name Ictalurus punctatus #common name channel catfish
DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
ACCESSIONS 20-Mar-1998
REFERENCE A05166; A05167
AUTHORS Andrews, P.C.; Ronner, P.
#journal J. Biol. Chem. (1985) 260:3910-3914
#title Isolation and structures of glucagon and glucagon-like
        peptide from catfish pancreas.
#cross-references MUID:85157536
#accession A05166
##molecule_type protein
##residues 1-29 ##label AND1
#accession A05167
##molecule_type protein
##residues 30-63 ##label AND2
CLASSIFICATION superfamily glucagon
KEYWORDS carbohydrate metabolism; duplication; hormone; pancreas
FEATURE
1-29      #product glucagon #status experimental #label GCN\
30-63     #product glucagon-like peptide 1 #status experimental
          #label GL1
SUMMARY   #length 63 #checksum 9366

Query Match       55.1%; Score 125; DB 1; Length 63;
Best Local Similarity 48.4%; Pred. No. 5.97e-08;
Matches 15; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Db 30 HADGTVTSYDSSYLQDAAKDFITWLSKGQP 60
|::|||::|::|::|::|::|::|::|::|
QY 1 hgegtfslskmqeeavrlfiewlknngp 31

RESULT 6

```

#journal	Gen. Comp. Endocrinol. (1991) 82:23-32
#title	The primary structure of glucagon-like peptide but not insulin has been conserved between the American eel, Anguilla rostrata and the European eel, Anguilla anguilla.
#cross-references	MUID:91340068
#accession	B61125
#molecule_type	protein
#residues	1-30 #label CON
CLASSIFICATION	#superfamily glucagon
FEATURE	amidated carboxyl end; duplication
1-1-30	
30	#product glucagon-like peptide #status experimental
	#label GLP\
	#modified_site amidated carboxyl end (Arg) #status predicted
SUMMARY	#length 30 #molecular-weight 3376 #checksum 6092
Query Match	53.3%; Score 121; DB 2; Length 30;
Best Local Similarity	48.3%; Pred. No. 2.93e-07;
Matches	14; Conservative 9; Mismatches 6; Indels 0; Gaps 0;
Db	1 HAEGTYSVSSYLQDAAKEFVSNLKTG 29
	: : : : : : : : : :
QY	1 hgegtfslskgmeeeavrlfiewlknk 29
	: : : : : : : : : :
RESULT	9
ENTRY	C61125 #type complete
TITLE	glucagon-like peptide - European eel
ORGANISM	#formal_name Anguilla anguilla #common_name European eel
DATE	10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 21-Nov-1997
ACCESSIONS	C61125
REFERENCE	A61125
#authors	Conlon, J.M.; Andrews, P.C.; Thim, L.; Moon, T.W.
#journal	Gen. Comp. Endocrinol. (1991) 82:23-32
#title	The primary structure of glucagon-like peptide but not insulin has been conserved between the American eel, Anguilla rostrata and the European eel, Anguilla anguilla
#cross-references	MUID:91340068
#accession	C61125
#molecule_type	protein
#residues	1-30 #label CON
CLASSIFICATION	#superfamily glucagon
KEYWORDS	amidated carboxyl end; duplication
FEATURE	#product glucagon-like peptide #status experimental
1-30	#label GLP\
30	#modified_site amidated carboxyl end (Arg) #status experimental
SUMMARY	#length 30 #molecular-weight 3376 #checksum 6092
Query Match	53.3%; Score 121; DB 2; Length 30;
Best Local Similarity	48.3%; Pred. No. 2.93e-07;
Matches	14; Conservative 9; Mismatches 6; Indels 0; Gaps 0;
Db	1 HAEGTYSVSSYLQDAAKEFVSNLKTG 29
	: : : : : : : : : :
QY	1 hgegtfslskgmeeeavrlfiewlknk 29
	: : : : : : : : : :
RESULT	10
ENTRY	I51093 #type fragment
TITLE	glucagon - chinook salmon (fragment)
ORGANISM	#formal_name Oncorhynchus tshawytscha #common_name chinook salmon
DATE	13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Nov-1997
ACCESSIONS	I51093
REFERENCE	A55895
#authors	Irwin, D.M.; Wong, J.
#journal	Mol. Endocrinol. (1995) 9:267-277
#title	Trout and chicken proglucagon: alternative splicing generates

```

#accession S06339
#molecule_type protein
#residues 1-36 #label POL
#accession S06871
#molecule_type protein
#residues 39-72 #label PO2
COMMENT X's at positions 37-38 represent a pair of basic amino acids
forming a cleavage site.
#superfamily glucagon
CLASSIFICATION
KEYWORDS carbohydrate metabolism; duplication; hormone; pancreas
FEATURES

```

Search completed: Mon Oct 4 15:23:58 1999
Job time : 8 secs.

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Oct 4 15:22:56 1999; MasPar time 4.19 Seconds
Molecular output not generated. 209.327 Million cell updates/sec

File: >MOHAM-312-CLAIM82A.PEP
Description: (1-31) from moham312177.pep
Sequence: 1 hgegtftdskmqeaeavrlfiewlknngp 31

Scoring table: PAM 150
Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot37
1:swissprot

Statistics: Mean 34.930; Variance 58.252; scale 0.600

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	227	100.0	87	1	EXE4_HEL SU	1.24e-30
2	220	96.9	39	1	EXE3_HEL SU	4.43e-29
3	127	55.9	103	1	GLUC_RANCA	1.45e-09
4	126	55.5	71	1	GLUC_ICTPU	2.28e-09
5	121	53.3	30	1	GLUC_ANGAN	2.07e-08
6	113	52.4	78	1	GLUC_LEPSP	4.95e-08
7	119	52.4	121	1	GLUC_CARAU	4.95e-08
8	118	52.0	122	1	GLU2_LOPAM	7.64e-08
9	117	51.5	68	1	GLUC_ONCKI	1.18e-07
10	115	50.7	29	1	GLUC_SCVCA	2.79e-07
11	115	50.7	158	1	GLUC_PIG	2.79e-07
12	115	50.7	180	1	GLUC_MOUSE	2.79e-07
13	115	50.7	180	1	GLUC_HUMAN	2.79e-07
14	115	50.7	180	1	GLUC_CAVPO	2.79e-07
15	115	50.7	180	1	GLUC_OCTDE	2.79e-07
16	115	50.7	180	1	GLUC_MESAU	2.79e-07
17	115	50.7	180	1	GLUC_RAT	2.79e-07
18	115	50.7	180	1	GLUC_BOVIN	2.79e-07
19	114	50.2	151	1	GLUC_CHICK	4.28e-07
20	113	49.8	124	1	GLU1_LOPAM	6.56e-07
21	110	48.5	29	1	GLUC_CHIBR	2.34e-06
22	108	47.6	29	1	GLUC_TORMA	5.44e-06
23	105	46.3	33	1	GLU2_ORENI	1.90e-05

24	104	45.8	29	1	GLUC_PLAPE	GLUCAGON.	2.86e-05
25	104	45.8	96	1	GLUC_MTOSC	GLUCAGON. PRECURSOR (FR	2.86e-05
26	100	44.1	29	1	GLUC_RABIT	GLUCAGON.	1.47e-04
27	100	44.1	69	1	GLUC_CANFA	GLUCAGON. PRECURSOR (FR	1.47e-04
28	99	43.6	29	1	GLUC_CALMI	GLUCAGON.	2.20e-04
29	99	43.6	75	1	GLUC_AMICA	GLUCAGON. PRECURSOR (FR	2.20e-04
30	98	43.2	36	1	GLU1_ORENI	GLUCAGON I.	3.29e-04
31	96	42.3	29	1	GLUC_DIDMA	GLUCAGON.	7.30e-04
32	95	41.9	29	1	GLUC_ANAPL	GLUCAGON.	1.08e-03
33	91	40.1	36	1	GLUC_HYDCO	GLUCAGON.	5.18e-03
34	88	38.8	2127	1	RRPL_RABVS	RNA POLYMERASE BETA SU	1.63e-02
35	88	38.8	2142	1	RRPL_RABVP	RNA POLYMERASE BETA SU	1.63e-02
36	83	36.6	406	1	YBDN_ECOLI	HYPOTHETICAL 47.8 KD P	1.05e-01
37	81	35.7	658	1	UVRB_HELPY	EXONUCLEASE ABC SUBUN	2.18e-01
38	80	35.2	170	1	VIP_HUMAN	VASOACTIVE INTESTINAL	3.13e-01
39	79	34.8	42	1	GIP_BOVIN	GASTRIC INHIBITORY POL	4.47e-01
40	79	34.8	42	1	GIP_PIG	GASTRIC INHIBITORY POL	4.47e-01
41	79	34.8	144	1	GIP_MOUSE	GASTRIC INHIBITORY POL	4.47e-01
42	79	34.8	144	1	GIP_RAT	GASTRIC INHIBITORY POL	4.47e-01
43	79	34.8	153	1	GIP_HUMAN	GASTRIC INHIBITORY POL	4.47e-01
44	79	34.8	328	1	NTPC_ENTHR	V-TYPE SODIUM ATP SYNT	4.47e-01
45	78	34.4	134	1	SECR_RAT	SECRETIN PRECURSOR.	6.37e-01

ALIGNMENTS

RESULT 1
ID EXE4_HEL SU STANDARD; PRG; 87 AA.
AC P26349;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE EXENDIN-4 PRECURSOR.
OS HELODERMA SUSPECTUM (GILA MONSTER).
OC EUFARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;
OC SCLEROGLOSSA; ANGUIMORPHA; HELODERMATIDAE; HELODERMA.
[1]
RP SEQUENCE FROM N.A.
RA CHEN Y.E., DRUCKER D.J.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 48-86.
RC TISSUE=VENOM;
RX MEDLINE; 92218391.
RA ENG J., KLEINMAN W.A., SINGH L., SINGH G., RAUFMAN J.-P.;
RT "Isolation and characterization of exendin-4, an exendin-3 analogue,
from Heloderma suspectum venom. Further evidence for an exendin
receptor on dispersed acini from guinea pig pancreas.";
RL J. BIOL. CHEM. 267:7402-7405(1992).
CC -!- FUNCTION: HAS A VIP/SECRETIN-LIKE BIOLOGICAL ACTIVITY. INTERACTS
WITH THE EXENDIN RECEPTOR.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; U77613; G1916067; -.
DR FIR; A42486; HWGH4G.
DR PROSITE; PS00260; GLUCAGON; 1.
DR PAM; PF00123; hormone2; 1.
KW GLUCAGON FAMILY; VENOM; AMIDATION; SIGNAL.
FT SIGNAL 1 23
FT PEPTIDE 48 86
FT MOD_RES 86
SQ SEQUENCE 87 AA; 9479 MW; 6C1A8FD5 CRC32;
Query Match 100.0%; Score 227; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.24e-30;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 48 HGGTFTSDLSKQMEEAARVLFIEWLKNKGGP 78
QY 1 hgegtftsdlksqmeeeavrlfiewlknkngp 31

RESULT 2

ID EXE3_HELHO STANDARD; PRT; 39 AA.
AC P20394;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE EXENDIN-3.
OS HELODERMA HORRIDUM HORRIDUM (MEXICAN BEADED LIZARD).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;
SCLEROGLOSSA; ANGUIMORPHA; HELODERMATIDAE; HELODERMA.
[1]

SEQUENCE

RC TISSUE-VENOM;
RX MEDLINE; 91056067.
RA ENG J., ANDREW P.C., KLEINMAN W.A., SINGH L., RAUFMAN J.-P.;
RT "Purification and structure of exendin-3, a new pancreatic
secretagogue isolated from Heloderma horridum venom."
RL J. BIOL. CHEM. 265:20259-20262(1990).
CC -!- FUNCTION: HAS A VIP/SECRETIN-LIKE BIOLOGICAL ACTIVITY. INTERACTS
CC WITH THE EXENDIN RECEPTOR.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC PIR; A23674; HWGH3Z.
DR PROSITE; PS00260; GLUCAGON; 1.
DR PFAM; PF00123; hormone2; 1.
DR HSSP; P01274; IGCN.
KW GLUCAGON FAMILY; VENOM; AMIDATION.
FT MOD-RES 39
SQ SEQUENCE 39 AA; 4204 MW; AB598FD3 CRC32;

Query Match 96.9%; Score 220; DB 1; Length 39;
Best Local Similarity 93.5%; Pred. No. 4.43e-29;
Matches 29; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 HSDGFTSDLSKQMEEAARVLFIEWLKNKGGP 31
QY 1 hgegtftsdlksqmeeeavrlfiewlknkngp 31

ILT 3

GLUC-RANCA STANDARD; PRT; 103 AA.
AC P15439; P15439; P15440;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE GLUCAGON PRECURSOR (FRAGMENTS).
OS RANA CATESBEIANA (BULL FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC NEOBATRACHIA; RANOIDEA; RANIDAE; RANINAE; RANA.
[1]

SEQUENCE

RC TISSUE-PANCREAS;
RX MEDLINE; 88257102.
RA POLLOCK H.G., HAMILTON J.W., ROUSE J.B., EBERN K.E., RAWITCH A.B.;
RT "Isolation of peptide hormones from the pancreas of the bullfrog
(Rana catesbeiana). Amino acid sequences of pancreatic polypeptide,
oxyntomodulin, and two glucagon-like peptides."
RL J. BIOL. CHEM. 263:9746-9751(1988).
CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC THE BLOOD SUGAR LEVEL.
CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC -!- Y'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH OTHER SPECIES
CC SEQUENCES.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR PIR; B28091; GCFGB.
DR PROSITE; PS00260; GLUCAGON; 3.

PFAM; PF00123; hormone2; 3.
DR HSSP; P01274; IGCN.
KW GLUCAGON FAMILY; HORMONE.
FT PEPTIDE 1 29
FT PEPTIDE 1 36
FT PEPTIDE 39 70
FT NON-CONS 70 71
FT PEPTIDE 71 103
SQ SEQUENCE 103 AA; 11719 MW; D43EDFC9 CRC32;

Query Match 55.9%; Score 127; DB 1; Length 103;
Best Local Similarity 51.6%; Pred. No. 1.46e-09;
Matches 16; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 39 HADGFTSDMSYLEEKAKEFVDMWLKGR 69
QY 1 hgegtftsdlksqmeeeavrlfiewlknkngp 31

RESULT 4

ID GLUC-ICTPU STANDARD; PRT; 71 AA.
AC P04093;
DT 01-NOV-1986 (REL. 03, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
DE GLUCAGON PRECURSOR (FRAGMENT).
OS ICTALURUS PUNCTATUS (CHANNEL CATFISH).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; SILURIFORMES; ICTALURIDAE;
OC ICTALURUS.
RN SEQUENCE.
RC TISSUE-PANCREAS;
RX MEDLINE; 87156787.
RA HOSEIN N.M., MAHREHOLZ A.M., ANDREWS P.C., GURD R.S.;
RT "Biological activities of catfish glucagon and glucagon-like
peptide."
RL BIOCHEM. BIOPHYS. RES. COMMUN. 143:87-92(1987).
RN SEQUENCE.
RC TISSUE-PANCREAS;
RX MEDLINE; 85157536.
RA ANDREWS P.C., RONNER P.;
RT "Isolation and structures of glucagon and glucagon-like peptide from
catfish pancreas."
RL J. BIOL. CHEM. 260:3910-3914(1985).
CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC -!- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH AMERICAN
CC GOOSEFISH SEQUENCES.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC PIR; A05166; GCIDC.
DR PROSITE; PS00260; GLUCAGON; 2.
DR PFAM; PF00123; hormone2; 2.
DR HSSP; P01274; IGCN.
KW GLUCAGON FAMILY; HORMONE.
FT NON-TER 1 1
FT PEPTIDE 1 29
FT PEPTIDE 38 71
FT CONFLICT 53 53
FT NON-TER 71 71
SQ SEQUENCE 71 AA; 8173 MW; C49ED93A CRC32;

Query Match 55.5%; Score 126; DB 1; Length 71;
Best Local Similarity 51.6%; Pred. No. 2.28e-09;
Matches 16; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Db 38 HADGFTSDMSYLEEKAKEFVDMWLKSGP 68
QY 1 hgegtftsdlksqmeeeavrlfiewlknkngp 31

```

CC      -!- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLGY WITH AMERICAN
CC      GOOSEFISH SEQUENCES.
CC      -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC      PIR; S06339; GCGXA.
DR      PROSITE; PS00260; GLUCAGON; 2.
DR      PFAM; PF00123; hormone2; 2.
DR      HSSP; P01274; IGCN.
DR      GLUCAGON FAMILY; HORMONE.
KW      PEPTIDE      1 29
FT      PEPTIDE      1 36
FT      PEPTIDE      45 78
SQ      SEQUENCE    78 AA; 8990 MW; 509ED9D3 CRC32;

Query Match      52.4%; Score 119; DB 1; Length 78;
Best Local Similarity 44.8%; Pred. No. 4.95e-08;
Matches 13; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Db      45 HADGTYTSDVSYLQQAQKKFTVTLKQG 73
      ::::||||: ::::|:|:|:|
QY      1 hgeftfslskmqeeavrlfiewlknng 29

RESULT 7
ID      GLUC_CARAU      STANDARD;      PRT;      121 AA.
AC      F79695;
DT      01-NOV-1997 (REL. 35, CREATED)
DT      01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT      01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE      GLUCAGON PRECURSOR.
OS      CARASSIUS AURATUS (GOLDFISH).
OS      EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC      TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; CYPRINIFORMES; CYPRINOIDEA;
OC      CYPRINIDAE; CYPRININAE; CARASSIUS.
RN      [1]
RP      SEQUENCE FROM N.A.
RA      YOEN T.T.H., MOK P.Y., CHOW B.K.C.;
RL      SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RC      -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC      THE BLOOD SUGAR LEVEL.
CC      -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U65528; G176277; -.
DR      PROSITE; PS00260; GLUCAGON; 2.
DR      PFAM; PF00123; hormone2; 2.
DR      HSSP; P01274; IGCN.
DR      GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
KW      GLUCAGON FAMILY; HORMONE.
FT      SIGNAL      1 21
FT      PEPTIDE     22 47
FT      PEPTIDE     50 78
FT      PEPTIDE     88 121
SQ      SEQUENCE    121 AA; 13527 MW; DDB862CE CRC32;

Query Match      52.4%; Score 119; DB 1; Length 121;
Best Local Similarity 48.4%; Pred. No. 4.95e-08;
Matches 15; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Db      88 HAEGTYTSDISSFLRQQAQNFVAVLKSGQP 118
      :||||:||||:|:|:|:|:|:|:|
QY      1 hgeftfslskmqeeavrlfiewlknngp 31

RESULT 8
ID      GLU2_LOPAM      STANDARD;      PRT;      122 AA.
AC      AC      P04092;
DT      01-NOV-1986 (REL. 03, CREATED)

```

DT 01-NOV-1986 (REL. 03, LAST SEQUENCE UPDATE)
 DE GLUCAGON II PRECURSOR.
 OS LOPHIUS AMERICANUS (AMERICAN GOOSEFISH) (ANGLERFISH).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 OC TELEOSTEI; EUTELEOSTEI; PARACANTHOPTERYGII; LOPHIIFORMES; LOPHIIDAE;
 OC LOPHIUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 83135785.
 RA LUND P.K., GOODMAN R.H., MONTMINY M.R., DEE P.C., HABENER J.F.;
 RT "Anglerfish islet pre-proglucagon II. Nucleotide and corresponding
 RT amino acid sequence of the cDNA."
 RJ J. BIOL. CHEM. 258:3280-3284(1983).
 RN [2]
 RP PROCESSING.
 RX MEDLINE: 86286913.
 RA NOE B.D., ANDREWS P.C.;
 RT "Specific glucagon-related peptides isolated from anglerfish islets
 RT are metabolic cleavage products of (pre)proglucagon-II."
 RJ PEPTIDES 7:331-339(1986).
 CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: V00632; G64022; -.
 DR EMBL: J00933; G213353; -.
 DR PIR: A05150; GCAF2.
 DR PROSITE: PS00260; GLUCAGON; 2.
 DR PFAM: PF00123; hormone2; 2.
 DR HSSP: P01274; IGCN.
 DR GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
 FT SIGNAL 1 21
 FT PEPTIDE 22 49 GRPP (GLICENTINE RELATED POLYPEPTIDE).
 FT PEPTIDE 52 80 GLUCAGON II.
 FT PEPTIDE 89 119 GLUCAGON-LIKE PEPTIDE II.
 FT SEQUENCE 122 AA; 14171 MW; DFE63061 CRC32;
 Query Match 52.0%; Score 118; DB 1; Length 122;
 Best Local Similarity 44.8%; Pred. No. 7.84e-08;
 Matches 13; Conservative 10; Mismatches 6; Indels 0; Gaps 0;
 Db 89 HADGIVTSVSVYLOQAAKDFVSWLKAG 117
 QY 1 hgegtftsldskmeeeavrlfiewlkn 29
 RESULT 9
 ID GLUC_ONCKI STANDARD; PRT; 68 AA.
 AC P07449;
 DT 01-APR-1988 (REL. 07, CREATED)
 DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
 DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
 DE GLUCAGON PRECURSOR (FRAGMENT).
 OS ONCORHYNCHUS KISUTCH (COHO SALMON).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES;
 OC SALMONIDAE; ONCORHYNCHUS.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE: 86234328.
 RA PLISETSKAYA E., POLLOCK H.G., ROUSE J.B., HAMILTON J.W., KIMMEL J.R.,
 RA GOREMAN A.;

RT "Isolation and structures of coho salmon (Oncorhynchus kisutch)
 RT glucagon and glucagon-like peptide."
 RJ REGUL. PEPT. 14:57-67(1986).
 CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -!- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH AMERICAN
 CC GOOSEFISH SEQUENCES.
 CC -!- GLN-14 IS A UNIQUE SUBSTITUTION FROM LEUCINE IN OTHER KNOWN
 CC GLUCAGON SEQUENCES AND GLUCAGON-LIKE PEPTIDES.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PIR: JF0103; GCONC.
 DR PROSITE: PS00260; GLUCAGON; 2.
 DR PFAM: PF00123; hormone2; 2.
 DR HSSP: P01274; IGCN.
 DR GLUCAGON FAMILY; HORMONE.
 KW NON_TER 1 1
 FT PEPTIDE 1 29 GLUCAGON.
 FT PEPTIDE 38 68 GLUCAGON-LIKE PEPTIDE.
 FT NON_TER 68 68
 SQ SEQUENCE 68 AA; 7810 MW; 402B55D1 CRC32;
 Query Match 51.5%; Score 117; DB 1; Length 68;
 Best Local Similarity 41.4%; Pred. No. 1.18e-07;
 Matches 12; Conservative 12; Mismatches 5; Indels 0; Gaps 0;
 Db 38 HADGIVTSVSVYLOQAAKDFVSWLKSG 66
 QY 1 hgegtftsldskmeeeavrlfiewlkn 29
 RESULT 10
 ID GLUC_SCYCA STANDARD; PRT; 29 AA.
 AC P09687;
 DT 01-MAR-1989 (REL. 10, CREATED)
 DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
 DT 01-JAN-1990 (REL. 13, LAST ANNOTATION UPDATE)
 DE GLUCAGON.
 OS SCYLIORHINUS CANICULA (SPOTTED DOGFISH) (SPOTTED CATSHARK).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; CHONDRICTHYES;
 OC ELASMOBRANCHII; CARCHARINIFORMES; SCYLIORHINIDAE; SCYLIORHINUS.
 RN [1]
 RP SEQUENCE.
 RX TISSUE-PANCREAS;
 RX MEDLINE: 87190953.
 RA CONLON J.M., O'TOOLE L., THIM L.;
 RT "Primary structure of glucagon from the gut of the common dogfish
 RT (Scyliorhinus canicula)."
 RJ FEBS LETT. 214:50-56(1987).
 CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PIR: A26992; GCDF.
 DR PROSITE: PS00260; GLUCAGON; 1.
 DR PFAM: PF00123; hormone2; 1.
 DR HSSP: P01274; IGCN.
 KW GLUCAGON FAMILY; HORMONE.
 SQ SEQUENCE 29 AA; 3529 MW; 8CFE41FB CRC32;
 Query Match 50.7%; Score 115; DB 1; Length 29;
 Best Local Similarity 53.6%; Pred. No. 2.79e-07;
 Matches 15; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
 Db 1 HSEGTFTSDYSKYMNRRAKDFVQWLNM 28
 QY 1 hgegtftsldskmeeeavrlfiewlkn 28
 RESULT 11
 ID GLUC_PIG STANDARD; PRT; 158 AA.

AC P01274;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE GLUCAGON PRECURSOR (FRAGMENT).
 GN GCG.
 OS SUS SCROFA (PIG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE; 81248172.
 RA THIM L., MOODY A.J.;
 RT "The primary structure of porcine glicentin (proglucagon).";
 RL REGUL. PEPT. 2:139-150(1981).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE; 82221776.
 RA THIM L., MOODY A.J.;
 RT "The amino acid sequence of porcine glicentin.";
 RL PEPTIDES 2 SUPPL. 2:37-39(1981).
 RN [3]
 RP SEQUENCE OF 33-61.
 RA BROMER W.W., SINN L.G., BEHRENS O.K.;
 RT "The amino acid sequence of glucagon. V. Location of amide groups,
 acid degradation studies and summary of sequential evidence.";
 RL J. AM. CHEM. SOC. 79:2807-2810(1957).
 RN [4]
 RP SEQUENCE OF 78-107.
 RA ORSKOV C., BERSANI M., JOHNSEN A.H., HOEJRP P., HOLST J.J.;
 RT "Complete sequences of glucagon-like peptide-1 from human and pig
 small intestine.";
 RL J. BIOL. CHEM. 264:12826-12829(1989).
 RN [5]
 RP SEQUENCE OF 111-158.
 RA BUHL T., THIM L., KOFOED H., ORSKOV C., HARLING H., HOLST J.J.;
 RT "Naturally occurring products of proglucagon 111-160 in the porcine
 and human small intestine.";
 RL J. BIOL. CHEM. 263:8621-8624(1988).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RA SASAKI K., DOCKERILL S., ADAMIAK D.A., TICKLE I.J., BLUNDELL T.L.;
 RT "X-ray analysis of glucagon and its relationship to receptor
 binding.";
 RL NATURE 257:751-757(1975).
 CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 THE BLOOD SUGAR LEVEL.
 CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -1- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH HUMAN
 SEQUENCE.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PR; A01540; GCGP.
 DR PDB; 1GCM; 30-SEP-83.
 DR PROSITE; PS00260; GLUCAGON; 3.
 DR PFAM; PF00123; hormone2; 3.
 KW GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES;
 3D-STRUCTURE.
 FT NON_TER 1 1
 FT PEPTIDE 1 30 GRPP (GLICENTINE RELATED POLYPEPTIDE).
 FT PEPTIDE 33 61 GLUCAGON.
 FT PEPTIDE 78 107 GLUCAGON-LIKE PEPTIDE 1.
 FT PEPTIDE 126 158 GLUCAGON-LIKE PEPTIDE 2.
 FT HELIX 39 42
 FT TURN 43 45
 FT TURN 46 55
 FT TURN 56 57
 SQ SEQUENCE 158 AA; 18212 MW; 9FBC1BFE CRC32;

Query Match

50.7%; Score 115; DB 1; Length 158;

Best Local Similarity 55.2%; Pred. No. 2.79e-07;
 Matches 16; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
 Db 78 HAEGTTSVSYLGGQAAKEFTAWLVKG 106
 QY 1 hgegttsdlskqmeeeavrlfiewlknk 29
 GCG.
 RESULT 12
 ID GLUC_MOUSE STANDARD; PRT; 180 AA.
 AC P55095;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE GLUCAGON PRECURSOR.
 GN GCG.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PANCREATIC ISLETS;
 RX MEDLINE; 95247722.
 RA ROTHENBERG M.E., EILERTSON C.D., KLEIN K., ZHOU Y., LINBERG I.,
 RA McDONALD J.K., MACKIN R.B., NOE B.D.;
 RT "Processing of mouse proglucagon by recombinant prohormone convertase
 1 and immunopurified prohormone convertase 2 in vitro.";
 RL J. BIOL. CHEM. 270:10136-10146(1995).
 CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 THE BLOOD SUGAR LEVEL.
 CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z46845; G599881; -
 CC MGD; MGI:95674; GCG.
 DR PROSITE; PS00260; GLUCAGON; 4.
 DR PFAM; PF00123; hormone2; 3.
 DR HSP; P01274; ICGN.
 KW GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
 FT SIGNAL 1 20 BY SIMILARITY.
 FT PEPTIDE 21 50 GRPP (GLICENTINE RELATED POLYPEPTIDE).
 FT PEPTIDE 53 81 GLUCAGON.
 FT PEPTIDE 92 128 GLUCAGON-LIKE PEPTIDE 1.
 FT PEPTIDE 146 178 GLUCAGON-LIKE PEPTIDE 2.
 SQ SEQUENCE 180 AA; 20906 MW; 0B21B7BA CRC32; .
 Query Match 50.7%; Score 115; DB 1; Length 180;
 Best Local Similarity 55.2%; Pred. No. 2.79e-07;
 Matches 16; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
 Db 98 HAEGTTSVSYLGGQAAKEFTAWLVKG 126
 QY 1 hgegttsdlskqmeeeavrlfiewlknk 29
 GCG.
 RESULT 13
 ID GLUC_HUMAN STANDARD; PRT; 180 AA.
 AC P01275;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE GLUCAGON PRECURSOR.
 GN GCG.
 OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE; 86330860.
 DR DRUCKER D.J., ASA S.;
 "Glucagon gene expression in vertebrate brain.";
 J. BIOL. CHEM. 263:13475-13478(1988).
 [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE; 86259053.
 KA WHITE J.W., SAUNDERS G.F.;
 "Structure of the human glucagon gene.";
 NUCLEIC ACIDS RES. 14:4719-4730(1986).
 [3]
 RN SEQUENCE FROM N.A.
 TX TISSUE=LIVER;
 MEDLINE; 83271477;
 BELL G.I., SANCHEZ-PESCADOR R., LAYBOURN P.J., NAJARIAN R.C.;
 "Exon duplication and divergence in the human preproglucagon gene.";
 NATURE 304:368-371(1983).
 [4]
 RN SEQUENCE OF 53-81.
 RA THOMSEN J., KRISTIANSEN K., BRUNFELDT K., SUNDBY F.;
 "The amino acid sequence of human glucagon.";
 FEBS LETT. 21:315-319(1972).
 [5]
 RN SEQUENCE OF 98-127.
 RX MEDLINE; 89327238.
 RA ORSKOV C., BERSANI M., JOHNSON A.H., HOEJURUP P., HOLST J.J.;
 "Complete sequences of glucagon-like peptide-1 from human and pig
 small intestine.";
 J. BIOL. CHEM. 264:12826-12829(1989).
 [6]
 RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 53-81.
 RX MEDLINE; 98334683.
 RA STURM N.S., LIN Y., BURLEY S.K., KRISTENANSKY J.L., AHN J.M.,
 AZIZEH B.Y., TRIVEDI D., HRUBY V.J.;
 "Structure-function studies on positions 17, 18, and 21 replacement
 analogues of glucagon: the importance of charged residues and salt
 bridges in glucagon biological activity.";
 J. MED. CHEM. 41:2693-2700(1998).
 CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 THE BLOOD SUGAR LEVEL.
 CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; J04040; G183270;
 DR EMBL; X03991; G762941;
 DR EMBL; V01515; G31778;
 DR EMBL; V01515; E28349; ALT_SEQ.
 DR PIR; A24377; GCHU.
 DR MIM; S23309; S23309.
 DR MIM; 138030;
 DR MIM; 231530;
 DR PROSITE; PS00260; GLUCAGON; 4.
 DR PFAM; PF00123; hormone2; 3.
 DR PDB; 1BHO; 18-NOV-98.
 KW GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL;
 3D-STRUCTURE. 1 20
 SIGNAL
 FT PEPTIDE 21 50 GRPP (GLICENTINE RELATED POLYPEPTIDE).
 FT PEPTIDE 53 81 GLUCAGON.
 FT PEPTIDE 98 127 GLUCAGON-LIKE PEPTIDE 1.

FT PEPTIDE 146 178 GLUCAGON-LIKE PEPTIDE 2.
 FT CONFLICT 82 K -> N (IN REF 3).
 SQ SEQUENCE 180 AA; 20909 MW; DEE43985 CRC32;
 Query Match 50.7%; Score 115; DB 1; Length 180;
 Best Local Similarity 55.2%; Pred. No. 2.79e-07;
 Matches 16; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
 Db 98 HAEGTFTSDVSYLGEQAKETIAVLK 126
 QY 1 hgegtftsdiskmeeavrlfiwllkng 29
 RESULT 14
 ID GLUC_CAVPO STANDARD; PRT; 180 AA.
 AC P05110;
 DT 13-AUG-1987 (REL. 05, CREATED)
 DT 01-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
 DT 13-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE GLUCAGON PRECURSOR.
 GN GCG.
 OS CAVIA PORCELLUS (GUINEA PIG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; HYSTRICOGNATHI; CAVIIDAE; CAVIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86248118.
 RA SEINO S., WELSH M., BELL G.I., CHAN S.J., STEINER D.F.;
 "Mutations in the guinea pig preproglucagon gene are restricted to a
 specific portion of the prohormone sequence.";
 FEBS LETT. 203:25-30(1986).
 RN [2]
 RP SEQUENCE OF 53-81.
 RX MEDLINE; 86165412.
 RA HUANG C.G., ENG J., PAN Y.-C.E., HULMES J.D., YALOW R.S.;
 "Guinea pig glucagon differs from other mammalian glucagons.";
 DIABETES 35:508-512(1986).
 RN [3]
 RP PARTIAL SEQUENCE OF 53-89.
 RX MEDLINE; 86017849.
 RA CONLON J.M., HANSEN H.F., SCHWARTZ T.W.;
 "Primary structure of glucagon and a partial sequence of
 oxyntomodulin (glucagon-37) from the guinea pig.";
 REGUL. PEPT. 11:309-320(1985).
 RL -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 THE BLOOD SUGAR LEVEL.
 CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D00014; D1000436;
 DR PIR; A24856; GCGP.
 DR PROSITE; PS00260; GLUCAGON; 4.
 DR PFAM; PF00123; hormone2; 3.
 DR HSP; P01274; IGCN.
 KW GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
 SIGNAL
 FT PEPTIDE 1 20
 FT PEPTIDE 21 50 GRPP (GLICENTINE RELATED POLYPEPTIDE).
 FT PEPTIDE 53 81 GLUCAGON.
 FT PEPTIDE 92 128 GLUCAGON-LIKE PEPTIDE 1.
 FT PEPTIDE 146 178 GLUCAGON-LIKE PEPTIDE 2.
 SQ SEQUENCE 180 AA; 20972 MW; 98724097 CRC32;
 Query Match 50.7%; Score 115; DB 1; Length 180;

Search completed: Mon Oct 4 15:23:02 1999
Job time : 6 secs.

(TM)

Alt No.	Score	Query Match	Length	ID	Description	Pred. No.
1	154	67.8	266	042143	PROGLUCAGON I.	1.68e-13
2	151	66.5	219	042144	PROGLUCAGON II.	6.15e-13
3	120	52.9	66	091188	GLUCAGON (FRAGMENT)	2.67e-07
4	120	52.9	72	0911409	PROGLUCAGON (FRAGMENT)	2.67e-07
5	120	52.9	72	0911408	PROGLUCAGON (FRAGMENT)	2.67e-07
6	120	52.9	178	091371	GLUCAGON I.	2.67e-07
7	120	52.9	178	091189	GLUCAGON II.	2.67e-07
8	114	50.2	206	091410	PROGLUCAGON.	2.94e-06
9	108	47.6	149	012955	PROGLUCAGON.	3.10e-05
10	108	47.6	204	012956	PROGLUCAGON.	3.10e-05
11	95	41.9	2127	057294	L PROTEIN, RNA DEPENDENT REPLICASE.	4.25e-03
12	86	37.9	552	9	HYPOTHETICAL 42.3 KD P	1.08e-01
13	83	37.9	2	085863	ROHS5.4 PROTEIN.	1.66e+00
14	77	33.9	414	021964	300A LONG HYPOTHETICAL	2.31e+00
15	76	33.5	300	059151	FORMATE DEHYDROGENASE	3.20e+00
16	76	33.5	1015	067146	NICOTINAMIDE NUCLEOTID	3.20e+00
17	76	33.5	1086	013423	NAD(P) TRANSHYDROGENAS	3.20e+00
18	76	33.5	1086	016796	NICOTINAMIDE NUCLEOTID	3.20e+00
19	76	33.5	1086	061941	C6.1A.	3.20e+00
20	75	33.0	324	016107		4.43e+00

RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

BEST LOCAL SIMILARITY 01.56, RECD NO. 1.000 157
 Matches 19: Conservative 7: Mismatches 5: Indels 0: Gaps 0:

1 bgeatf+sd\skameeeavr\few\knagp 37

042144: AC

DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)

00 MESOBATRACHIA: PIPOIDEA; PIPIIDAE; XENOPUS; RANA; AMPHIBIENS; ANUROIDEA; ANURIDAE; ANURA

RN [1]

```

DR PFAM; PF00123; hormone2; 2.  
FT NON_TER      1  
SQ SEQUENCE     72 AA;   8293 MW;    OF7AF3EC CRC32;  
  
Query Match          52.9%; Score 120; DB 13; Length 72;  
Best Local Similarity 44.8%; Pred.No. 2.67e-07;  
Matches 13; Conservative 11; Mismatches 5; Indels 0; Gaps 0;  
  
Ddb       39 HADGTYTSDVSTYLQQAADFVSWLKSG 67  
         |::|:||||: :||: ||: |:|:-  
OY        1 hgegtfstdlskqmeeeavrlfielwkg 29  
         |::|:||||: :||: ||: |:|:-  
  
RESULT           5  
ID Q91408 PRELIMINARY; PRT; 72 AA.  
AC Q91408;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE PROGLUCAGON [FRAGMENT].  
OS ONCORHYNCHUS MYXISS (RAINBOW TROUT) (SALMO GAIRDNERI).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;  
OC TELEOSTEI; EULELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;  
ONCORHYNCHUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 95295739.  
RA IRWIN D.M., WONG J.;  
RT "Trout and chicken proglucagon: alternative splicing generates mRNA  
transcripts encoding glucagon-like peptide 2.";  
RL MOL. ENDOCRINOL. 9:267-277(1995).  
DR EMBL; S78473; G999383; -.  
DR PFAM; PF00123; hormone2; 2.  
FT NON_TER      1  
FT NON_TER      72  
SQ SEQUENCE     72 AA;   8293 MW;    OF7AF3EC CRC32;  
  
Query Match          52.9%; Score 120; DB 13; Length 72;  
Best Local Similarity 44.8%; Pred.No. 2.67e-07;  
Matches 13; Conservative 11; Mismatches 5; Indels 0; Gaps 0;  
  
Ddb       39 HADGTYTSDVSTYLQQAADFVSWLKSG 67  
         |::|:||||: ~~~~: |::~|:  
OY        1 hgegtfstdlskqmeeeavrlfielwkg 29  
         |::|:||||: ~~~~: |::~|:  
  
RESULT           6  
ID Q91971 PRELIMINARY; PRT; 178 AA.  
AC Q91971;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE GLUCAGON I.  
OS ONCORHYNCHUS MYXISS (RAINBOW TROUT) (SALMO GAIRDNERI).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;  
OC TELEOSTEI; EULELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;  
ONCORHYNCHUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-INTESTINE, DISTAL PORTION;  
RX MEDLINE; 95295739.  
RA IRWIN D.M., WONG J.;  
RT "Trout and chicken proglucagon: alternative splicing generates mRNA  
transcripts encoding glucagon-like peptide 2.";  
RL MOL. ENDOCRINOL. 9:267-277(1995).  
DR EMBL; U19917; G736385; -.  
DR EMBL; S78475; G999385; -.  
DR PROSITE; PS00260; GLUCAGON; 3.  
DR PFAM; PF00123; hormone2; 3.  
SQ SEQUENCE     178 AA;  20034 MW;  2056F963 CRC32;  
  
Query Match          52.9%; Score 120; DB 13; Length 178;  
Best Local Similarity 44.8%; Pred.No. 2.67e-07;
```

```
Matches 13; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

Db 90 HADGTYSDVSTYLODQAAKDFVSWLKS 118
   |||||:||||:||||:|:|||||
QY 1 hgefttsdskqmeeeavrlfiewlkg 29

RESULT 7
ID Q91189 PRELIMINARY; PRT; 178 AA.
AC Q91189; Q92158;
DT 01-NOV-1996 (TREMELREL. 01, CREATED)
DT 01-NOV-1996 (TREMELREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMELREL. 08, LAST ANNOTATION UPDATE)
DE GLUCAGON II.
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIRDNERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;
ON ONCORHYNCHUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE, DIGITAL PORTION;
RC MEDLINE: 95295739.
IRWIN D.M., WONG J.;
"Trout and chicken proglucagon: alternative splicing generates mRNA
transcripts encoding glucagon-like peptide 2.,"
MOL. ENDOCRINOL. 9:267-277(1995).
DR EMBL; U19914; G736363; -.
DR EMBL; U19916; G736372; -.
DR EMBL; U19915; G736372; JOINED.
DR EMBL; U19915; G736371; -.
DR PFAM; PF00123; hormone2; 3.
SQ SEQUENCE 178 AA; 19998 MW; A4299C13 CRC32;

Query Match 52.98; Score 120; DB 13; Length 178;
Best Local Similarity 44.88; Pred. No. 2.67e-07;
Matches 13; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

Db 90 HADGTYSDVSTYLODQAAKDFVSWLKS 118
   |||||:||||:||||:|:|||||
QY 1 hgefttsdskqmeeeavrlfiewlkg 29

RESULT 8
ID Q91410 PRELIMINARY; PRT; 206 AA.
AC Q91410;
DT 01-NOV-1996 (TREMELREL. 01, CREATED)
DT 01-NOV-1996 (TREMELREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMELREL. 08, LAST ANNOTATION UPDATE)
DE PROGLUCAGON.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95295739.
IRWIN D.M., WONG J.;
"Trout and chicken proglucagon: alternative splicing generates mRNA
transcripts encoding glucagon-like peptide 2.,"
MOL. ENDOCRINOL. 9:267-277(1995).
DR EMBL; S78477; G999387; -.
DR PROSITE; PS00260; GLUCAGON; 3.
DR PFAM; PF00123; hormone2; 3.
SQ SEQUENCE 206 AA; 23875 MW; 8EC91118 CRC32;

Query Match 50.24; Score 114; DB 13; Length 206;
Best Local Similarity 51.74; Pred. No. 2.94e-06;
Matches 15; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 116 HADGRTSDISSYLEGQAAKEFTIAVLNG 146
   |||||:||||:|:|:|:|:|:|:|
QY 1 hgefttsdskqmeeeavrlfiewlkg 29

Query Match 47.68; Score 108; DB 13; Length 204;
Best Local Similarity 48.38; Pred. No. 3.10e-05;
Matches 14; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
```

```
RESULT 9
ID Q12955 PRELIMINARY; PRT; 149 AA.
AC Q12955;
DT 01-JUL-1997 (TREMELREL. 04, CREATED)
DT 01-JUL-1997 (TREMELREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMELREL. 08, LAST ANNOTATION UPDATE)
DE PROGLUCAGON.
GN LPI.
OS HELODERMA SUSPECTUM (GILA MONSTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SOUAMATA;
OC SCLEROGLOSSA; ANGUIMORPHA; HELODERMATIDAE; HELODERMA.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN Y.E., DROCKER D.J.;
RL J. BIOL. CHEM. 0:0-0(0).
DR EMBL; U77611; G1916063; -.
DR PROSITE; PS00260; GLUCAGON; 1.
DR PFAM; PF00123; hormone2; 2.
SQ SEQUENCE 149 AA; 17224 MW; F763AB51 CRC32;

Query Match 47.68; Score 108; DB 13; Length 149;
Best Local Similarity 48.38; Pred. No. 3.10e-05;
Matches 14; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 116 HADGRTSDISSYLEGQAAKEFTIAVLNG 144
   |||||:||||:|:|:|:|:|:|:|
QY 1 hgefttsdskqmeeeavrlfiewlkg 29

RESULT 10
ID Q12956 PRELIMINARY; PRT; 204 AA.
AC Q12956;
DT 01-JUL-1997 (TREMELREL. 04, CREATED)
DT 01-JUL-1997 (TREMELREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMELREL. 08, LAST ANNOTATION UPDATE)
DE PROGLUCAGON.
GN LPII.
OS HELODERMA SUSPECTUM (GILA MONSTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SOUAMATA;
OC SCLEROGLOSSA; ANGUIMORPHA; HELODERMATIDAE; HELODERMA.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN Y.E., DROCKER D.J.;
RL J. BIOL. CHEM. 0:0-0(0).
DR EMBL; U77612; G1916065; -.
DR PROSITE; PS00260; GLUCAGON; 2.
DR PFAM; PF00123; hormone2; 3.
SQ SEQUENCE 204 AA; 23553 MW; EE50250D CRC32;

Query Match 47.68; Score 108; DB 13; Length 204;
Best Local Similarity 48.38; Pred. No. 3.10e-05;
Matches 14; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 116 HADGRTSDISSYLEGQAAKEFTIAVLNG 144
   |||||:||||:|:|:|:|:|:|:|
QY 1 hgefttsdskqmeeeavrlfiewlkg 29

RESULT 11
ID Q57294 PRELIMINARY; PRT; 2127 AA.
AC Q57294;
DT 01-JUN-1998 (TREMELREL. 06, CREATED)
DT 01-JUN-1998 (TREMELREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMELREL. 06, LAST ANNOTATION UPDATE)
DE L PROTEIN, RNA DEPENDENT RNA POLYMERASE.
GN L.
OS RABIES VIRUS.
OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; MONONEGAVIRALES; RHABDOVIRIDAE;
OC LYSSAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RC-HL;
RA MINAMOTO N.;
```


RT "Complete Sequence and Gene Organization of the Genome of a
Hyper-thermophilic Archaeobacterium, Pyrococcus horikoshii OT3.";

RL DNA RES. 5:55-76(1998)

DR EMEL; AP000006; D1031532; -.

DR PROSITE; PS00782; TFIIB; 2.

KW INITIATION FACTOR.

SO SEQUENCE 300 AA; 34097 MW; 6E17BB64 CRC32;

Query Match

Best Local Similarity 33.5%; Score 76; DB 1; Length 300;

Matches 10; Conservativity 50.0%; Pred.No. 3.20e+00;

Mismatches 7; Indels 3; Gaps 0;

Db 125 LPKHVEEAEARLYREAVRKG 144

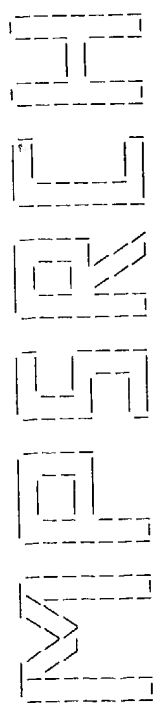
|:::|||||::|::|

OY 10 lskqneeeavrlfiewlkng 29

Search completed: Mon Oct 4 15:23:33 1999

Job time : 14 secs.

Mohamed, A.
09/312177
segs. Claims 82-84



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Oct 4 15:28:18 1999; Maspar time 4.70 Seconds
140.296 Million cell updates/sec

ular output not generated.

Title: >MOHAM-312-CLAIM84.PEP
Description: (1-31) from moham312177.pep
Perfect Score: 220
Sequence: 1 diskqmeeeavrlfiewlknngpssgappps 31

Claim 84

Scoring table:
PAM 150
Gap 11

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:
a-geneseq35
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Statistics: Mean 24.023; Variance 96.248; scale 0.250
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	220	100.0	31	14	R80547 Heloderma suspectum e	1.10e-12
2	220	100.0	39	39	W61769 Extendin-3, for use in	1.10e-12
3	220	100.0	39	39	W61770 Extendin-4, for use in	1.10e-12
4	220	100.0	39	30	W47608 Gila monster extendin-	1.10e-12
5	220	100.0	39	14	R80545 Heloderma horridum ex	1.10e-12
6	220	100.0	39	14	R80546 Heloderma suspectum e	1.10e-12
7	220	100.0	39	30	W47609 Gila monster extendin-	1.10e-12
8	220	100.0	87	35	W70288 Heloderma suspectum p	1.10e-12
9	195	88.6	39	39	W61773 Leu(14), Phe(25)-exen	4.22e-08
10	173	78.6	30	14	R80543 Heloderma suspectum e	2.47e-07
11	165	75.0	30	39	W61771 Extendin-4 (1-30) for	2.47e-07
12	165	75.0	31	14	R80544 Heloderma suspectum e	1.15e-06
13	158	71.8	30	29	W39368 H. horridum extendin-3	1.15e-06
14	158	71.8	30	29	W39301 H. horridum extendin-4	1.15e-06
15	158	71.8	30	29	W39302 H. horridum extendin-4	1.15e-06
16	153	69.5	30	29	W39309 H. horridum extendin-4	3.42e-06

17	151	68.6	28	39	W61772	Extendin-4 (1-28) amid	5.28e-06
18	149	67.7	30	29	W39339	H. horridum extendin-3	8.15e-06
19	149	67.7	30	29	W39312	H. horridum extendin-4	8.15e-06
20	147	66.8	28	29	W39375	H. horridum extendin-3	1.26e-05
21	147	66.8	30	29	W39385	H. horridum extendin-4	1.26e-05
22	147	66.8	30	29	W39332	H. horridum extendin-3	1.26e-05
23	147	66.8	30	29	W39379	H. horridum extendin-4	1.26e-05
24	147	66.8	30	29	W39382	H. horridum extendin-3	1.26e-05
25	147	66.8	30	29	W39334	H. horridum extendin-4	1.26e-05
26	147	66.8	30	29	W39304	H. horridum extendin-3	1.26e-05
27	147	66.8	30	29	W39319	H. horridum extendin-4	1.26e-05
28	147	66.8	30	29	W39395	H. horridum extendin-3	1.26e-05
29	147	66.8	30	29	W39386	H. horridum extendin-4	1.26e-05
30	147	66.8	30	29	W39318	H. horridum extendin-3	1.26e-05
31	147	66.8	30	29	W39370	H. horridum extendin-4	1.26e-05
32	147	66.8	30	29	W39370	H. horridum extendin-3	1.26e-05
33	147	66.8	30	29	W39325	H. horridum extendin-4	1.26e-05
34	147	66.8	30	29	W39420	H. horridum extendin-3	1.26e-05
35	147	66.8	30	29	W39308	H. horridum extendin-4	1.26e-05
36	147	66.8	30	29	W39333	H. horridum extendin-3	1.26e-05
37	147	66.8	30	29	W39330	H. horridum extendin-4	1.26e-05
38	147	66.8	30	29	W39335	H. horridum extendin-3	1.26e-05
39	147	66.8	30	29	W39388	H. horridum extendin-4	1.26e-05
40	147	66.8	30	29	W39329	H. horridum extendin-3	1.26e-05
41	147	66.8	30	29	W39331	H. horridum extendin-4	1.26e-05
42	147	66.8	30	29	W39394	H. horridum extendin-3	1.26e-05
43	147	66.8	30	29	W39390	H. horridum extendin-4	1.26e-05
44	147	66.8	30	29	W39327	H. horridum extendin-3	1.26e-05
45	147	66.8	30	29	W39380	H. horridum extendin-4	1.26e-05

ALIGNMENTS

RESULT 1
ID R80547 standard; peptide; 31 AA.
AC R80547;1996 (first entry)
DE Heloderma suspectum extendin-4 residues 9-39 (Extendin-4(9-39)).
KW Extendin-4; residues 9-39; Extendin-4(9-39);
OS insulintropic peptides; inhibitor.
OS Heloderma suspectum.
PN US5424286-A.
PD 13-JUN-1995.
PF 24-MAY-1993; 066480.
PR 24-MAY-1993; US-066480.
PA (ENGJ/) ENG J.
PI Eng J;
PI WPI; 95-262627/34.
DR Stimulating/inhibiting insulin release with extendin polypeptide(s) -
PT for treating diabetes mellitus and preventing hyperglycaemia.
PS Claim 7; Columns 13-14; 17pp; English.
CC R80547 is the Heloderma suspectum extendin-4 residues 9-39. It
is an insulintropic peptide activity inhibitor.
SQ Sequence 31 AA;
Query Match 100.0%; Score 220; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.10e-12;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 diskqmeeeavrlfiewlknngpssgappps 31
QY 1 diskqmeeeavrlfiewlknngpssgappps 31
RESULT 2
ID W61769 standard; peptide; 39 AA.
AC W61769;
DE Extendin-3, for use in treating disorders related to food intake.
KW Extendin-3; obesity; type II diabetes; eating disorders; cardiac disease;
OS insulin resistance syndrome; elevated plasma glucose level; agonist.
OS Heloderma horridum.
PN WO9830231-A1.

PD 16-JUL-1998. U00449.
 PF 07-JAN-1998; US-066029.
 PR 14-NOV-1997; US-066029.
 PR 07-JAN-1997; US-034905.
 PR 08-AUG-1997; US-055404.
 PR 14-NOV-1997; US-065442.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Beley NRA, Bhavsar S, Prickett KS;
 DR WPI: 98-398796/34.
 DT Reducing food intake by administering extendins or their
 PT analogues - for treatment of e.g. obesity, type II diabetes,
 PT eating disorders and insulin resistance
 PS Claims 16, 24; Page 8; 214pp; English.
 CC The invention relates to a new method for treating disorders that
 CC are alleviated by reducing food intake, in particular obesity, type
 CC II diabetes, eating disorders, insulin resistance syndrome, elevated
 CC plasma glucose levels, or the risk of cardiac disease. The method
 CC comprises administering an extendin or an extendin agonist. The treatment
 CC reduces appetite and lowers plasma lipid levels. It inhibits food
 CC consumption as effectively as amylin or cholecystokinin but has a much
 CC longer-lasting action (still effective after 6 hours in a mouse model).
 CC The present sequence is that of extendin-3 which is one of the preferred
 CC compounds for use in the method.
 SQ Sequence 39 AA;

Query Match 100.0%; Score 220; DB 39; Length 39;
 Best Local Similarity 100.0%; Pred. No. 1.10e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 dlskmeeeavrflfiewlknpgssgappps 39
 |||||
 QY 1 dlskmeeeavrflfiewlknpgssgappps 31

RESULT 3
 ID W61770 standard; peptide; 39 AA.
 DT 29-MAR-1999 (first entry)
 DE Extendin-4, for use in treating disorders related to food intake.
 KW Extendin; obesity; type II diabetes; eating disorders; cardiac disease;
 KW insulin resistance syndrome; elevated plasma glucose level; agonist.
 OS Heloderma suspectum.
 PN WO9830231-A1.
 PL 16-JUL-1998.
 DT 07-JAN-1998; U00449.
 DT 14-NOV-1997; US-066029.
 DT 07-JAN-1997; US-034905.
 DT 08-AUG-1997; US-055404.
 DT 14-NOV-1997; US-065442.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Beley NRA, Bhavsar S, Prickett KS;
 DR WPI: 98-398796/34.
 DT Reducing food intake by administering extendins or their
 PT analogues - for treatment of e.g. obesity, type II diabetes,
 PT eating disorders and insulin resistance
 PS Claims 17, 25; Page 8; 214pp; English.
 CC The invention relates to a new method for treating disorders that
 CC are alleviated by reducing food intake, in particular obesity, type
 CC II diabetes, eating disorders, insulin resistance syndrome, elevated
 CC plasma glucose levels, or the risk of cardiac disease. The method
 CC comprises administering an extendin or an extendin agonist. The treatment
 CC reduces appetite and lowers plasma lipid levels. It inhibits food
 CC consumption as effectively as amylin or cholecystokinin but has a much
 CC longer-lasting action (still effective after 6 hours in a mouse model).
 CC The present sequence is that of extendin-4 which is one of the preferred
 CC compounds for use in the method.
 SQ Sequence 39 AA;

Query Match 100.0%; Score 220; DB 39; Length 39;
 Best Local Similarity 100.0%; Pred. No. 1.10e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 dlskmeeeavrflfiewlknpgssgappps 39

QY 1 dlskmeeeavrflfiewlknpgssgappps 31
 |||||

RESULT 4
 ID W47608 standard; peptide; 39 AA.
 AC W47608;
 DT 03-JUL-1998 (first entry)
 DE Gila monster extendin-3
 KW Extendin agonist; gastric motility; gastric emptying; treatment;
 KW spasms; postprandial dumping syndrome; postprandial hyperglycaemia;
 KW type 1 diabetes; impaired glucose tolerance; toxin ingestion;
 KW obesity; Gila monster venom; extendin-3.
 OS Heloderma horridum.
 PH Key Location/Qualifiers
 FT Modified_site 39 /note="amidated"
 FT WO9805351-A1.
 PN 12-FEB-1998.
 PD 08-AUG-1997; U14199.
 PF 08-AUG-1996; US-694954.
 PR (AMYL-) AMYLIN PHARM INC.
 PA Beley NRA, Gedin B, Prickett KS, Young AA;
 PI WPI: 98-145351/13.
 DR Regulating gastrointestinal motility using extendins or their
 PT agonists - for treating spasms, diabetic postprandial hyperglycaemia,
 PT impaired glucose tolerance etc., also in diagnostic investigations
 PS Claims 20 and 21; Fig 1: 70pp; English.
 CC W47549 describes a generic extendin agonist, provided that it does
 CC have the formula of either extendin-3 (W47608) or extendin-4
 CC (W47609).

CC Extendin agonists, which reduce gastric motility and delay gastric
 CC emptying, can be used to treat spasm (where associated with acute
 CC diverticulitis or disorders of the biliary tract or sphincter of
 CC Oddi), postprandial dumping syndrome and hyperglycaemia
 CC (particularly associated with type 2 diabetes), type 1 diabetes, is
 CC impaired glucose tolerance, toxin ingestion (an extendin agonist is
 CC administered to prevent stomach contents passing into the
 CC intestines, then the stomach pumped) and obesity. They can also be
 CC administered to subjects undergoing gastrointestinal diagnostic
 CC investigation, particularly radiological or by magnetic resonance
 CC imaging.
 CC Extendins, components of Gila monster venom, have some sequence
 CC similarity to glucagon-like peptides (GLP). They are GLP agonists
 CC and have been suggested (US5424286) for treatment of diabetes and
 CC prevention of hyperglycaemia.
 SQ Sequence 39 AA;

Query Match 100.0%; Score 220; DB 30; Length 39;
 Best Local Similarity 100.0%; Pred. No. 1.10e-12;
 Matches -31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 dlskmeeeavrflfiewlknpgssgappps 39
 |||||
 QY 1 dlskmeeeavrflfiewlknpgssgappps 31

RESULT 5
 ID R80545 standard; peptide; 39 AA.
 AC R80545;
 DT 27-FEB-1996 (first entry)
 DE Heloderma horridum extendin-3.
 KW Extendin-3; diabetes mellitus; hyperglycaemia; insulinotropic peptide.
 OS Heloderma horridum.
 PN US5424286-A.
 PD 13-JUN-1995.
 PF 24-MAY-1993; 066480.
 PR 24-MAY-1993; US-066480.
 PA (ENGJ/) ENG J.
 DR WPI: 95-262627/34.
 PT Stimulating/inhibiting insulin release with extendin polypeptide(s) -
 PT for treating diabetes mellitus and preventing hyperglycaemia.


```

PS Claim 5; Columns 13-14; 17pp; English.
CC R80545 is Heloderma horridum exendin-3. It is an
CC insulinotropic peptide, and can therefore be used in the treatment of
CC diabetes mellitus (types I or II), and for the prevention of
CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
CC and insulin-(in)dependent mechanisms.
SQ Sequence 39 AA;

Query Match 100.0%; Score 220; DB 14; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.10e-12;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 dlskmeeeavrlfiwknkgpssgappps 39
Qy 1 dlskmeeeavrlfiwknkgpssgappps 31

RESULT 6
ID R80546 standard; peptide; 39 AA.
AC R80546;
DE 27-FEB-1996 (first entry)
HE Heloderma suspectum exendin-4.
EX Exendin-4; diabetes mellitus; hyperglycaemia; insulinotropic peptide.
OS Heloderma suspectum.
PD US5424286-A.
PN 13-JUN-1995.
PF 24-MAY-1993; 066480.
PR 24-MAY-1993; US-066480.
PA (ENGJ) ENG J.
PI Eng J;
DR WPI; 95-262627/34.
PT Stimulating/inhibiting insulin release with exendin polypeptide(s) -
PT for treating diabetes mellitus and preventing hyperglycaemia.
PS Claim 6; Columns 13-14; 17pp; English.
CC R80546 is Heloderma suspectum exendin-4. It is an
CC insulinotropic peptide, and can therefore be used in the treatment of
CC diabetes mellitus (types I or II), and for the prevention of
CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
CC and insulin-(in)dependent mechanisms.
SQ Sequence 39 AA;

Query Match 100.0%; Score 220; DB 14; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.10e-12;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 dlskmeeeavrlfiwknkgpssgappps 39
Qy 1 dlskmeeeavrlfiwknkgpssgappps 31

RESULT 7
AC W47609 standard; peptide; 39 AA.
DE 03-JUL-1998 (first entry)
EX Exendin agonist; gastric motility; gastric emptying; treatment;
KW spasm; postprandial dumping syndrome; postprandial hyperglycaemia;
KW type 1 diabetes; impaired glucose tolerance; toxin ingestion;
KW obesity; Gila monster venom; exendin-4.
OS Heloderma suspectum.
FH Key Location/Qualifiers
FT Peptide 1..23 /note= "amidated"
FT Peptide 1..47 /note= "Signal peptide"
FT Peptide 1..47 /note= "ENTP"
FT Peptide 48..87 /note= "ENTP"
FT Peptide 48..87 /note= "Exendin 4"
FT Cleavage_site 46..47 /note= "Dipeptidyl peptidase cleavage site"
PN W09835033-A1.
PD 13-AUG-1998.
PF 04-FEB-1998; CA0071.
PR 07-FEB-1997; GB-002582.
PR 05-FEB-1997; US-037412.
PA (ONEO-) 1149336 ONTARIO INC.
PI Drucker DU;
DR WPI; 98-447230/38.
DR N-PSDB; V33163.
PT New nucleic acid encoding proexendin - used to diagnose and treat,
PT e.g. endocrine tumours, also to treat poisoning by reptile venom
PS Claim 3; Fig 2; 26pp; English.
CC The Heloderma suspectum proexendin peptide is encoded by its cDNA
CC which was isolated from a H. suspectum salivary gland cDNA library.
CC The proexendin protein comprises of a novel exendin N-terminal
CC peptide (ENTP) linked to the N-terminus of the exendin 4 peptide
CC by a consensus dipeptidyl peptidase cleavage site. The proexendin
CC cDNA can be used to clone or identify related sequences (e.g. the
CC exendin 3 gene of Heloderma horridum, mutant alleles and proexendin
CC gene regulatory defects associated with metabolic disease) and species
CC homologues (e.g. for developing animal models for drug screening).
CC The proexendin peptide can be used to raise antibodies. Anti-proexendin
CC antibodies are claimed to be useful for diagnosing conditions associated
CC with altered levels of proexendin (e.g. endocrine tumours and organ

```

```

PS Claims 20 and 21; Fig 1; 70pp; English.
CC W47549 describes a generic exendin agonist, provided that it does
CC have the formula of either exendin-3 (W47608) or exendin-4
CC (W47609).
CC Exendin agonists, which reduce gastric motility and delay gastric
CC emptying, can be used to treat spasm (where associated with acute
CC diverticulitis or disorders of the biliary tract or sphincter of
CC Oddi), postprandial dumping syndrome and hyperglycaemia
CC (particularly associated with type 2 diabetes), type 1 diabetes,
CC impaired glucose tolerance, toxin ingestion (an exendin agonist is
CC administered to prevent stomach contents passing into the
CC intestines, then the stomach pumped) and obesity. They can also be
CC administered to subjects undergoing gastrointestinal diagnostic
CC investigation, particularly radiological or by magnetic resonance
CC imaging.
CC Exendins, components of Gila monster venom, have some sequence
CC similarity to glucagon-like peptides (GLP). They are GLP agonists
CC and have been suggested (US5424286) for treatment of diabetes and
CC prevention of hyperglycaemia.
SQ Sequence 39 AA;

Query Match 100.0%; Score 220; DB 30; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.10e-12;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 dlskmeeeavrlfiwknkgpssgappps 39
Qy 1 dlskmeeeavrlfiwknkgpssgappps 31

RESULT 8
ID W70288 standard; Protein; 87 AA.
AC W70288;
DE 06-NOV-1998 (first entry)
HE Heloderma suspectum proexendin peptide.
KW Heloderma suspectum proexendin; exendin N-terminal peptide; ENTP;
KW exendin 4 peptide; exendin 3 gene; Heloderma horridum; metabolic disease;
KW drug screening; endocrine tumour; organ failure; cell metabolism;
KW diabetes; reptilian venom peptide.
OS Heloderma suspectum.
FH Key Location/Qualifiers
FT Peptide 1..23 /note= "Signal peptide"
FT Peptide 1..47 /note= "ENTP"
FT Peptide 48..87 /note= "ENTP"
FT Peptide 48..87 /note= "Exendin 4"
FT Cleavage_site 46..47 /note= "Dipeptidyl peptidase cleavage site"
PN W09835033-A1.
PD 13-AUG-1998.
PF 04-FEB-1998; CA0071.
PR 07-FEB-1997; GB-002582.
PR 05-FEB-1997; US-037412.
PA (ONEO-) 1149336 ONTARIO INC.
PI Drucker DU;
DR WPI; 98-447230/38.
DR N-PSDB; V33163.
PT New nucleic acid encoding proexendin - used to diagnose and treat,
PT e.g. endocrine tumours, also to treat poisoning by reptile venom
PS Claim 3; Fig 2; 26pp; English.
CC The Heloderma suspectum proexendin peptide is encoded by its cDNA
CC which was isolated from a H. suspectum salivary gland cDNA library.
CC The proexendin protein comprises of a novel exendin N-terminal
CC peptide (ENTP) linked to the N-terminus of the exendin 4 peptide
CC by a consensus dipeptidyl peptidase cleavage site. The proexendin
CC cDNA can be used to clone or identify related sequences (e.g. the
CC exendin 3 gene of Heloderma horridum, mutant alleles and proexendin
CC gene regulatory defects associated with metabolic disease) and species
CC homologues (e.g. for developing animal models for drug screening).
CC The proexendin peptide can be used to raise antibodies. Anti-proexendin
CC antibodies are claimed to be useful for diagnosing conditions associated
CC with altered levels of proexendin (e.g. endocrine tumours and organ

```


KW Exendin-4: residues 1-31; v-31-Exendin-4(1-31); diabetes mellitus;
 KW hyperglycaemia; Tyr31; insulinotropic peptide.
 OS Heloderma suspectum.

PN US5424286-A.

PD 13-JUN-1995.

PF 24-MAY-1993; 066480.

PR 24-MAY-1993; US-056480.

PA (ENG/J) ENG J.

Eng J;

PI WPI; 95-262627/34.

PT Stimulating/inhibiting insulin release with exendin polypeptide(s) -
 PT for treating diabetes mellitus and preventing hyperglycaemia.

PS Claim 2; Columns 13-14; 17pp; English.

CC R80544 is the Heloderma suspectum exendin-4 residues 1-31, where

CC the native Pro1 has been replaced with a Tyr residue. It is an

CC insulinotropic peptide, and can therefore be used in the treatment of

CC diabetes mellitus (types I or II), and for the prevention of

CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent

CC and insulin-(in)dependent mechanisms.

SQ Sequence 31 AA;

Query Match 75.0%; Score 165; DB 14; Length 31;

Best Local Similarity 100.0%; Pred. No. 2.47e-07;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 dlskqmeeeavrflfiewlknng 30

QY 1 dlskqmeeeavrflfiewlknng 22

RESULT 13

ID W39368 standard; peptide; 30 AA.

AC W39368;

DT 05-JUN-1998 (first entry)

DE H. horridum exendin-3 peptide derivative #11.

KW Exendin-3; exendin 4; diabetes; insulin; secretion; biosynthesis;

KW glucagon reduction; hypoglycaemia; glucose; treatment.

OS Heloderma horridum.

Key Location/Qualifiers

FT Modified_site 30 /note= "C-terminal amide"

W09746584-A1.

PI 11-DEC-1997.

PD 05-JUN-1997; E02930.

PF 13-SEP-1996; DE-037230.

PR 05-JUN-1996; DE-022502.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Goetze B, Goetze R, Hoffmann E;

WPI; 98-042119/04.

PT Truncated versions of exendin peptide(s) for treating diabetes -

PT increase secretion and biosynthesis of insulin, but reduce those of

PT glucagon, and do not induce hypoglycaemia

PS Claim 2; Page 27; 150pp; English.

CC Peptides W39303-W39420 are fragments of exendin-3 and exendin-4

CC isolated from Heloderma horridum which are used in a novel method

CC for the treatment of diabetes mellitus. These peptides can stimulate

CC biosynthesis and secretion of insulin, but have the opposite effect on

CC glucagon, and independent of this activity can increase peripheral

CC glucose utilisation. Exendin-3 and exendin-4 are only active when blood

CC sugar levels are high, so they will not induce hypoglycaemia. Compared

CC with glucagon-like peptide 1 (GLP1) and the known exendins, they are more

CC active (effective at lower doses), more stable to degradation and

CC metabolism and have a longer lasting effect. Truncated forms of this

CC peptide can be made more economically than full length versions.

SQ Sequence 30 AA;

Query Match 71.8%; Score 158; DB 29; Length 30;

Best Local Similarity 100.0%; Pred. No. 1.15e-06;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 dlskqmeeeavrflfiewlknng 29

QY 1 dlskqmeeeavrflfiewlknng 21

RESULT 14

ID W39301 standard; peptide; 30 AA.

AC W39301;

DT 05-JUN-1998 (first entry)

DE H. horridum exendin-3 peptide.

KW Exendin-3; exendin 4; diabetes; insulin; secretion; biosynthesis;

KW glucagon reduction; hypoglycaemia; glucose; treatment.

OS Heloderma horridum.

Key Location/Qualifiers

FT Modified_site 30

/note= "This residue can be any amino acid except

Gly"

W09746584-A1.

PI 11-DEC-1997.

PD 05-JUN-1997; E02930.

PF 13-SEP-1996; DE-037230.

PR 05-JUN-1996; DE-022502.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Goetze B, Goetze R, Hoffmann E;

WPI; 98-042119/04.

PT Truncated versions of exendin peptide(s) for treating diabetes -

PT increase secretion and biosynthesis of insulin, but reduce those of

PT glucagon, and do not induce hypoglycaemia

PS Claim 1; Page 3; 150pp; English.

CC This peptide is a fragment of exendin-3 isolated from Heloderma

CC horridum. This peptide and its salts, esters and derivatives can be

CC used to treat diabetes mellitus. They stimulate biosynthesis and

CC secretion of insulin, but have the opposite effect on glucagon, and

CC independent of this activity can increase peripheral glucose utilisation.

CC Exendin-3 and exendin-4 are only active when blood sugar levels are

CC high, so they will not induce hypoglycaemia. Compared with glucagon-like

CC peptide 1 (GLP1) and the known exendins, they are more active (effective

CC at lower doses), more stable to degradation and metabolism and have a

CC longer lasting effect. Truncated forms of this peptide can be made more

CC economically than full length versions.

SQ Sequence 30 AA;

Query Match 71.8%; Score 158; DB 29; Length 30;

Best Local Similarity 100.0%; Pred. No. 1.15e-06;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 dlskqmeeeavrflfiewlknng 29

QY 1 dlskqmeeeavrflfiewlknng 21

RESULT 15

ID W39302 standard; peptide; 30 AA.

AC W39302;

DT 05-JUN-1998 (first entry)

DE H. horridum exendin-4 peptide.

KW Exendin-3; exendin 4; diabetes; insulin; secretion; biosynthesis;

KW glucagon reduction; hypoglycaemia; glucose; treatment.

OS Heloderma horridum.

Key Location/Qualifiers

FT Modified_site 30

/note= "This residue can be any amino acid except

Gly"

W09746584-A1.

PI 11-DEC-1997.

PD 05-JUN-1997; E02930.

PF 13-SEP-1996; DE-037230.

PR 05-JUN-1996; DE-022502.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Goetze B, Goetze R, Hoffmann E;

WPI; 98-042119/04.

PT Truncated versions of exendin peptide(s) for treating diabetes -

PT increase secretion and biosynthesis of insulin, but reduce those of

PT glucagon, and do not induce hypoglycaemia

PS Claim 1; Page 4; 150pp; English.

SQ This peptide is a fragment of exendin-4 isolated from Heloderma

CC horridum. This peptide and its salts, esters and derivatives can be
CC used to treat diabetes mellitus. They stimulate biosynthesis and
CC secretion of insulin, but have the opposite effect on glucagon, and
CC independent of this activity can increase peripheral glucose utilisation.
CC Exendin-3 and exendin-4 are only active when blood sugar levels are
CC high, so they will not induce hypoglycaemia. Compared with glucagon-like
CC peptide 1 (GLP1) and the known exendins, they are more active (effective
CC at lower doses), more stable to degradation and metabolism and have a
CC longer lasting effect. Truncated forms of this peptide can be made more
CC economically than full length versions.
SQ Sequence 30 AA;

Query Match 71.8%; Score 158; DB 29; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.15e-06;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
-b 9 diskqmeeeavrlfiewlkng 29
/ 1 diskqmeeeavrlfiewlkng 21

Search completed: Mon Oct 4 15:28:38 1999
Job time : 20 secs.

WQRLH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Oct 4 15:27:49 1999; MasPar time 4.29 Seconds

ular output not generated. 289.449 Million cell updates/sec

File: >MOHAM-312-CLAIM84.PEP
Description: (1-31) from moham312177.pep
Perfect Score: 220
Sequence: 1 dlskmeeeavrlfiewlknngpssgappps 31

Scoring table: PAM 150
Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir60
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 33.166; Variance 59.781; scale 0.555

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	220	100.0	39	1 HWGH4G	exendin-4 - Gila mons	2.54e-27
2	220	100.0	39	1 HWGH3Z	exendin-3 - Mexican b	2.54e-27
3	91	41.4	31	2 S4472	glucagon G2 - North A	8.43e-03
4	88	40.0	2127	1 ZLVNSB	genome polypeptide -	2.53e-02
5	88	40.0	2142	1 ZLVNSV	genome polypeptide -	2.53e-02
6	86	39.1	31	2 S4471	glucagon G1 - North A	5.20e-02
7	86	39.1	552	2 S45978	replicase - phase PF7	5.20e-02
8	82	37.3	63	1 GCTDC	glucagon precursor -	2.14e-01
9	80	36.4	30	2 S4473	glucagon-like peptide	4.29e-01
10	80	36.4	101	1 GCFGB	glucagon precursor -	4.29e-01
11	77	35.0	30	2 B61125	glucagon-like peptide	1.19e+00
12	77	35.0	30	2 C61125	glucagon-like peptide	1.19e+00
13	77	35.0	66	2 I51093	glucagon - chinook sa	1.19e+00
14	77	35.0	178	2 I51058	glucagon I precursor	1.19e+00
15	77	35.0	178	2 I51057	glucagon II precursor	1.19e+00
16	76	34.5	72	1 GCGXA	glucagon precursor -	1.67e+00
17	76	34.5	300	2 E71023	probable transcriptio	1.67e+00
18	76	34.5	2185	2 S60200	acetyl-CoA carboxylas	1.67e+00
19	75	34.1	92	2 G64837	probable acylphosphat	2.32e+00
20	75	34.1	122	1 GCAF2	glucagon 2 precursor	2.32e+00
21	75	34.1	207	1 XUHMC	methylated-DNA--prote	2.32e+00
22	75	34.1	406	2 H64793	ybDn protein - Escher	2.32e+00
23	74	33.6	60	1 GCONC	glucagon precursor -	3.23e+00

transcription factor 4.48e+00
probable integrase - 4.48e+00
citrate (pro-3S)-lyas 4.48e+00
RNA-directed DNA poly 4.48e+00
neurofascin - chicken 4.48e+00
probable ATP-dependen 4.48e+00
hypothetical protein 4.48e+00
hypothetical protein 6.20e+00
type I restriction en 8.54e+00
protein-tyrosinephos 8.54e+00
glucagon precursor - 1.17e+01
proglucagon - chicken 1.17e+01
urea transport protei 1.17e+01
polyprotein - echovir 1.17e+01
esterase - Spirulina 1.61e+01
ntpp protein - Entero 1.61e+01
exonuclease ABC chal 1.61e+01
exonuclease ABC chal 1.61e+01
genome polypeptide - 1.61e+01
probable transcriptio 2.20e+01
conserved hypotheticala 2.20e+01
acetyl-CoA carboxylas 2.20e+01

ALIGNMENTS

RESULT 1
ENTRY HWGH4G #type complete
TITLE exendin-4 - Gila monster
ORGANISM #formal_name Heloderma suspectum #common_name Gila monster
DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 21-Nov-1997
ACCESSIONS A42486
REFERENCE A42486
#authors Eng, J.; Kleinman, W.A.; Singh, L.; Singh, G.; Raufman, J.P.
#journal J. Biol. Chem. (1992) 267:7402-7405
#title Isolation and characterization of exendin-4, an exendin-3 analogue, from Heloderma suspectum venom. Further evidence for an exendin receptor on dispersed acini from guinea pig pancreas.
#cross-references MIMD:92218391
#accession A42486
#molecule_type protein
#residues 1-39 #label ENG
COMMENT Exendin-4 does not stimulate amylase secretion by pancreatic acinar cells.
CLASSIFICATION #superfamily glucagon
KEYWORDS amidated carboxyl end; duplication; venom
FEATURE 39
#modified_site amidated carboxyl end (Ser) #status experimental
SUMMARY #length 39 #molecular-weight 4188 #checksum 9570
Query Match 100.0%; Score 220; DB 1; Length 39;
Best Local Similarity 100.0%; Pred.No. 2.54e-27;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 DLSQMEEEAVRLFIEWLKNNGPSSGAPPPS 39
QY 1 dlskmeeeavrlfiewlknngpssgappps 31
RESULT 2
ENTRY HWGH3Z #type complete
TITLE exendin-3 - Mexican beaded lizard
ORGANISM #formal_name Heloderma horridum #common_name Mexican beaded lizard
DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 21-Nov-1997
ACCESSIONS A23674
REFERENCE A23674
#authors Eng, J.; Andrews, P.C.; Kleinman, W.A.; Singh, L.; Raufman, J.P.

```

attenuated rabies virus SAD B19.
#cross-references MUID:90223994
#accession E34746
#molecule_type genomic RNA
#residues 1-2127 #label CON
#cross-references GB:M31046; NID:g333556; PID:g333561
GENETICS
#gene L
#superfamily rhabdovirus L protein
#nucleotidyltransferase; RNA binding; RNA biosynthesis;
#transmembrane protein
KEYWORDS
FEATURE
543-562
1965-1982
SUMMARY
#region RNA binding #status predicted\
#domain transmembrane #status predicted #label TM2
#length 2127 #molecular-weight 242977 #checksum 9107
Query Match 40.0%; Score 88; DB 1; Length 2127;
Best Local Similarity 42.9%; Pred. No. 2,53e-02;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0
Db 37 NLNSPLIEDPARLMLEWLTG 57
:|: : |:: || :||| |
QY 1 diskgmeeavrfiewlKng 21

RESULT 5
ENTRY
TITLE Z1VNPV #type complete
CONTAINS genome polyprotein - rabies virus (strain PV)
CONTAINS L protein
CONTAINS RNA-directed RNA polymerase (EC 2.7.7.48)
ORGANISM #formal_name rabies virus
DATE 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change
29-May-1998
ACCESSION A29248; E24887
REFERENCE A29248
#authors Tordo, N.; Poch, O.; Ermine, A.; Keith, G.; Rougeon, F.
#journal Virology (1988) 165:565-576
#title Completion of the rabies virus genome sequence determination
of highly conserved domains among the L (polymerase) proteins
of unsegmented negative-strand RNA viruses.
#cross-references MUID:88306248
#accession A29248
#molecule_type genomic RNA
#residues 1-2142 #label TOR
#cross-references GB:M21634; GB:M21634; NID:g333585; PID:g333590
REFERENCE A94100
#authors Tordo, N.; Poch, O.; Ermine, A.; Keith, G.; Rougeon, F.
#journal Proc. Natl. Acad. Sci. U.S.A. (1986) 83:3914-3918
#title Walking along the rabies genome: is the large G-L intergenic
region a remnant gene?
#cross-references MUID:86233343
#accession E24887
#molecule_type DNA
#residues 1-28 #label TO2
GENETICS
#gene L
#superfamily rhabdovirus L protein
#nucleotidyltransferase; RNA binding; RNA biosynthesis;
#transmembrane protein
KEYWORDS
FEATURE
543-562
1965-1982
SUMMARY
#region RNA binding #status predicted\
#domain transmembrane #status predicted #label TM2
#length 2142 #molecular-weight 244484 #checksum 433
Query Match 40.0%; Score 88; DB 1; Length 2142;
Best Local Similarity 42.9%; Pred. No. 2,53e-02;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
Db 37 NLNSPLIEDPARLMLEWLTG 57
:|: : |:: || :||| |
QY 1 diskgmeeavrfiewlKng 21

```

```

6
RESULT
ENTRY      S44471      #type complete
TITLE      glucagon GI - North American paddlefish (Polyodon spathula)
ORGANISM   #formal_name Polyodon spathula
DATE       18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change
20-Mar-1998
ACCESSIONS S44471
REFERENCE   S44457
#authors   Nguyen, T.M.; Mommensen, T.P.; Mims, S.M.; Conlon, J.M.
#journal   Biochem. J. (1994) 300:339-345
#title     Characterization of insulins and proglucagon-derived peptides
           from a phylogenetically ancient fish, the paddlefish
           (Polyodon spathula).
#accession S44471
#molecule_type protein
#residues  1-31 #label NGU
#experimental_source pancreas
CLASSIFICATION #superfamily glucagon
KEYWORDS      carbohydrate metabolism; duplication; hormone; pancreas
FEATURE       1-31
#product glucagon GI #status predicted #label MAT
#length 31 #molecular_weight 3751 #checksum 7808
Query Match 39.1%; Score 86; DB 2; Length 31;
Best Local Similarity 57.1%; Pred. No. 5.20e-02;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
Db 9 DYSKYLEKRAKEFVWLNKG 29
| | | : | | : | | | | |
QY 1 dlskmeeeavrlfiewlknkng 21

7
RESULT
ENTRY      S46978      #type complete
TITLE      replicase - phage PP7
ORGANISM   #formal_name phage PP7
DATE       13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
09-Sep-1997
ACCESSIONS S46978
REFERENCE   S46975
#authors   Olsthoorn, R.C.L.; Garde, G.; Dayhuff, T.; Atkins, J.F.; Van
           Duin, J.
#submission submitted to the EMBL Data Library, July 1994
#description Nucleotide sequence of a single-stranded RNA phage from
           Pseudomonas aeruginosa; kinship to coliphages and
           conservation of regulatory RNA structures.
#accession S46978
#status     preliminary
#molecule_type mRNA
#residues   1-552 #label OLS
#cross-references EMBL:X80191; NID:g517237; PID:g517241
JMMARY      #length 552 #molecular_weight 63300 #checksum 8424
Query Match 39.1%; Score 86; DB 2; Length 552;
Best Local Similarity 45.0%; Pred. No. 5.20e-02;
Matches 9; Conservative 9; Mismatches 0; Indels 2; Gaps 2;
Db 483 DISKRLDDE-VR-YVDWLRN 500
| | | : | | : | | | | |
QY 1 dlskmeeeavrlfiewlkn 20

8
RESULT
ENTRY      GCDC      #type fragments
TITLE      glucagon precursor - channel catfish (fragments)
ORGANISM   #formal_name Ictalurus punctatus #common_name channel catfish
DATE       31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
20-Mar-1998
ACCESSIONS A05166; A05167
REFERENCE   A92514
#authors   Andrews, P.C.; Ronner, P.
#journal   J. Biol. Chem. (1985) 260:3910-3914
#title     Isolation and structures of glucagon and glucagon-like
           peptide from catfish pancreas.
#cross-references MUID:85157536
#accession A05166
#molecule_type protein
#residues   1-29 #label AND1
#accession A05167
#molecule_type protein
#residues   30-63 #label AND2
CLASSIFICATION #superfamily glucagon
KEYWORDS      carbohydrate metabolism; duplication; hormone; pancreas
FEATURE       1-29
#product glucagon #status experimental #label GCN
#product glucagon-like peptide 1 #status experimental
#label GL1
SUMMARY      #length 63 #checksum 9366
Query Match 37.3%; Score 82; DB 1; Length 63;
Best Local Similarity 43.5%; Pred. No. 2.14e-01;
Matches 10; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
Db 38 DVSSVLODQAAKDFITWLSGQP 60
| | | : | | : | | | | |
QY 1 dlskmeeeavrlfiewlknkngpp 23

9
RESULT
ENTRY      S44473      #type complete
TITLE      glucagon-like peptide - North American paddlefish (Polyodon
           spathula)
ORGANISM   #formal_name Polyodon spathula
DATE       18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change
20-Mar-1998
ACCESSIONS S44473
REFERENCE   S44467
#authors   Nguyen, T.M.; Mommensen, T.P.; Mims, S.M.; Conlon, J.M.
#journal   Biochem. J. (1994) 300:339-345
#title     Characterization of insulins and proglucagon-derived peptides
           from a phylogenetically ancient fish, the paddlefish
           (Polyodon spathula).
#accession S44473
#molecule_type protein
#residues   1-30 #label NGU
CLASSIFICATION #superfamily glucagon
KEYWORDS      duplication; hormone; pancreas
FEATURE       1-30
#product glucagon-like peptide #status predicted #label
           MAT
SUMMARY      #length 30 #molecular_weight 3359 #checksum 5186
Query Match 36.4%; Score 80; DB 2; Length 30;
Best Local Similarity 56.3%; Pred. No. 4.29e-01;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Db 14 LQEQAAEDFISWLKKG 29
| | | : | | : | | | | |
QY 6 meeeavrlfiewlknkng 21

10
RESULT
ENTRY      GCFCB      #type fragments
TITLE      glucagon precursor - bullfrog (fragments)
ORGANISM   #formal_name Rana catesbeiana #common_name bullfrog
DATE       31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
20-Mar-1998
ACCESSIONS B28091; C28091; D28091
REFERENCE   A92730
#authors   Pollock, H.G.; Hamilton, J.W.; Rouse, J.B.; Ebner, K.E.;
           Ravitch, A.B. (1988) 263:9746-9751
#journal   J. Biol. Chem. (1988) 263:9746-9751
#title     Isolation of peptide hormones from the pancreas of the
```


bullfrog (*Rana catesbeiana*). Amino acid sequences of pancreatic polypeptide, oxyntomodulin, and two glucagon-like peptides.

```
#cross-references MUID:89257102
#accession B28091
##molecule_type protein
##residues 1-36 #label P02
#accession C28091
##molecule_type protein
##residues 37-68 #label P0L
#accession D28091
##molecule_type protein
##residues 69-101 #label P03
CLASSIFICATION #superfamily glucagon
KEYWORDS carbohydrate metabolism; duplication; hormone; pancreas
FEATURE
1-36
#product glucagon-36 (oxyntomodulin) #status
experimental #label G36\
#product glucagon #status predicted #label GCN\
#product glucagon-like peptide 1 #status experimental
#label GL1\
#product glucagon-like peptide 2 #status experimental
#label GL2
SUMMARY
#length 101 #checksum 9108
Query Match 36.4%; Score 80; DB 1; Length 101;
Best Local Similarity 43.5%; Pred. No. 4.29e+01;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
```

```
Db 45 DMSSYLEEAAKEFVWLKGRP 67
|:| :||| :| :||| :|
QY 1 diskmeeeeavrlfiewkngpp 23
```

```
RESULT 11
ENTRY B61125 #type complete
TITLE glucagon-like peptide - American eel
ORGANISM #formal_name Anguilla rostrata #common_name American eel
DATE 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change
21-Nov-1997
ACCESSION B61125
REFERENCE A61125
#authors Conlon, J.M.; Andrews, P.C.; Thim, L.; Moon, T.W.
#journal Gen. Comp. Endocrinol. (1991) 82:23-32
#title The primary structure of glucagon-like peptide but not
insulin has been conserved between the American eel,
Anguilla rostrata and the European eel, Anguilla anguilla.
#cross-references MUID:91340068
#accession B61125
##molecule_type protein
##residues 1-30 #label CON
CLASSIFICATION #superfamily glucagon
KEYWORDS amidated carboxyl end; duplication
FEATURE
1-30
#product glucagon-like peptide #status experimental
#label GLP\
#modified_site amidated carboxyl end (Arg) #status
predicted
SUMMARY
#length 30 #molecular-weight 3376 #checksum 6092
Query Match 35.0%; Score 77; DB 2; Length 30;
Best Local Similarity 38.1%; Pred. No. 1.19e+00;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
```

```
Db 9 DVSSYLQDQAAKEFVSWLKTG 29
|:| :||| :| :||| :|
QY 1 diskmeeeeavrlfiewkng 21
```

```
RESULT 12
ENTRY C61125 #type complete
TITLE glucagon-like peptide - European eel
ORGANISM #formal_name Anguilla anguilla #common_name European eel
```

```
DATE 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change
21-Nov-1997
ACCESSION C61125
REFERENCE A61125
#authors Conlon, J.M.; Andrews, P.C.; Thim, L.; Moon, T.W.
#journal Gen. Comp. Endocrinol. (1991) 82:23-32
#title The primary structure of glucagon-like peptide but not
insulin has been conserved between the American eel,
Anguilla rostrata and the European eel, Anguilla anguilla.
#cross-references MUID:91340068
#accession C61125
##molecule_type protein
##residues 1-30 #label CON
CLASSIFICATION #superfamily glucagon
KEYWORDS amidated carboxyl end; duplication
FEATURE
1-30
#product glucagon-like peptide #status experimental
#label GLP\
#modified_site amidated carboxyl end (Arg) #status
experimental
SUMMARY
#length 30 #molecular-weight 3376 #checksum 6092
Query Match 35.0%; Score 77; DB 2; Length 30;
Best Local Similarity 38.1%; Pred. No. 1.19e+00;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
```

```
Db 9 DVSSYLQDQAAKEFVSWLKTG 29
|:| :||| :| :||| :|
QY 1 diskmeeeeavrlfiewkng 21
```

```
RESULT 13
ENTRY I51093 #type fragment
TITLE glucagon - chinook salmon (fragment)
ORGANISM #formal_name Oncorhynchus tshawytscha #common_name chinook
salmon
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
21-Nov-1997
ACCESSION I51093
REFERENCE A55895
#authors Irwin, D.M.; Wong, J.
#journal Mol. Endocrinol. (1995) 9:267-277
#title Trout and chicken proglucagon: alternative splicing generates
mRNA transcripts encoding glucagon-like peptide 2.
#cross-references MUID:95295739
#accession I51093
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-66 #label IRW
##cross-references EMBL:U19920; NID:g736366; PID:g736367
CLASSIFICATION #superfamily glucagon
KEYWORDS duplication
SUMMARY
#length 66 #checksum 1440
```

```
Query Match 35.0%; Score 77; DB 2; Length 66;
Best Local Similarity 38.1%; Pred. No. 1.19e+00;
Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
```

```
Db 41 DVSTYLQDQAAKEFVSWLKS 61
|:| :||| :| :||| :|
QY 1 diskmeeeeavrlfiewkng 21
```

```
RESULT 14
ENTRY I51058 #type complete
TITLE glucagon I precursor - rainbow trout
ORGANISM #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
21-Nov-1997
ACCESSION I51058; I51299; I51056; I51037; I51036; I51300
REFERENCE A55895
#authors Irwin, D.M.; Wong, J.
#journal Mol. Endocrinol. (1995) 9:267-277
```

#title Trout and chicken proglucagon; alternative splicing generates
#cross-references EMBL:U19917; NID:g736364; PID:g736365; GB:S78475;
#accession I51058
##status preliminary; translated from GB/EMBL/DBDJ
##molecule_type mRNA
##residues 1-178 #label IRW
##cross-references EMBL:U19917; NID:g736364; PID:g736365; GB:S78475;
#accession I51056
##status preliminary; translated from GB/EMBL/DBDJ
##molecule_type mRNA
##residues 58-123 #label IR3
##cross-references EMBL:U19913; NID:g736360; PID:g736361
#accession I51037
##status preliminary; translated from GB/EMBL/DBDJ
##molecule_type DNA
##residues 'M',114-144 #label IR4
##cross-references EMBL:U19919; NID:g736374; PID:g736377
#accession I51036
##status preliminary; translated from GB/EMBL/DBDJ
##molecule_type DNA
##residues 113-123 #label IR5
##cross-references EMBL:U19918; NID:g736373; PID:g736376

GENETICS

#introns 123/2
CLASSIFICATION #superfamily glucagon
KEYWORDS duplication
SUMMARY #length 178 #molecular-weight 20034 #checksum 5250

Query Match 35.0%; Score 77; DB 2; Length 178;
Best Local Similarity 38.1%; Pred. No. 1.19e+00;
Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
LQ 98 DVSTYLDQQAQKDFVSWLXSG 118
|:| :|:|:| :|:|:|
QY 1 diskmeeeavrflfiewlknq 21

RESULT 15
ENTRY I51057 #type complete
TITLE glucagon II precursor - rainbow trout
ORGANISM #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Nov-1997
CROSS-REFERENCES I51057; I51039; I51038
REFERENCE R55895
#authors Irwin, D.M.; Wong, J.
#journal Mol. Endocrinol. (1995) 9:267-277
#title Trout and chicken proglucagon; alternative splicing generates
#cross-references EMBL:U19917; NID:g736364; PID:g736365; GB:S78475;
#accession I51058
##status preliminary; translated from GB/EMBL/DBDJ
##molecule_type mRNA
##residues 1-178 #label IRW
##cross-references EMBL:U19914; NID:g736362; PID:g736363
#accession I51039
##status preliminary; translated from GB/EMBL/DBDJ
##molecule_type DNA
##residues 113-144 #label IR2
##cross-references EMBL:U19916; NID:g736369; PID:g736372
#accession I51038
##status preliminary; translated from GB/EMBL/DBDJ
##molecule_type DNA
##residues 113-123 #label IR3
##cross-references EMBL:U19915; NID:g736368; PID:g736371

GENETICS

#introns 123/2
CLASSIFICATION #superfamily glucagon
KEYWORDS duplication
SUMMARY #length 178 #molecular-weight 19998 #checksum 4464
Query Match 35.0%; Score 77; DB 2; Length 178;
Best Local Similarity 38.1%; Pred. No. 1.19e+00;
Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Db 98 DVSTYLDQQAQKDFVSWLXSG 118
|:| :|:|:| :|:|:|
QY 1 diskmeeeavrflfiewlknq 21

Search completed: Mon Oct 4 15:28:01 1999
Job time : 12 secs.

MOHAM-312-CLAIM84.PEP

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Oct 4 15:26:51 1999; MasPar time 3.25 Seconds
ular output not generated.

>MOHAM-312-CLAIM84.PEP
Description: (1-31) from moham312177.pep
Perfect Score: 220
Sequence: 1 dlsgmeeeavrlfiewlknngpssgappps 31

Scoring table: PAM 150
Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot37
1:swissprot

Statistics: Mean 33.941; Variance 55.026; scale 0.617

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	220	100.0	39	1	EXE3_HELHO	1.73e-30
2	220	100.0	87	1	EXE4_HELHU	1.73e-30
3	88	40.0	2127	1	RRLP_RABVS	6.16e-03
4	88	40.0	2142	1	RRLP_RABVP	6.16e-03
5	83	37.7	71	1	GLUC_ICTPU	4.33e-02
6	80	36.4	103	1	GLUC_RANCA	1.35e-01
7	77	35.0	30	1	GLUC_RANGA	4.12e-01
8	77	35.0	33	1	GLUC_ORANI	4.12e-01
9	76	34.5	78	1	GLUC_LEPSP	5.93e-01
10	75	34.1	92	1	ACTP_ECOLI	8.51e-01
11	75	34.1	121	1	GLUC_CARAU	8.51e-01
12	75	34.1	122	1	GLU2_LOPAP	8.51e-01
13	75	34.1	207	1	MGMT_HUMAN	8.51e-01
14	75	34.1	389	1	UTL_HUMAN	8.51e-01
15	75	34.1	406	1	YBDN_ECOLI	8.51e-01
16	74	33.6	68	1	GLUC_ONCKI	1.22e+00
17	73	33.2	261	1	TF2B_PYRWO	1.74e+00
18	73	33.2	300	1	TF2B_PYRFO	1.74e+00
19	73	33.2	323	1	TRBB_AGRT6	1.74e+00
20	73	33.2	402	1	VINT_BPPH8	1.74e+00
21	73	33.2	510	1	CILA_ECOLI	1.74e+00
22	71	32.3	227	1	COX2_LOCOM	3.50e+00
23	71	32.3	1075	1	Y124_METJA	3.50e+00

24	70	31.8	151	1	GLUC_CHICK	GLUCAGON PRECURSOR.	4.94e+00
25	70	31.8	204	1	Y4DW_RHSN	HYPOTHETICAL 22.9 KD P	4.94e+00
26	69	31.4	328	1	NTPC_ENTHR	V-TYPE SODIUM ATP SYNT	6.94e+00
27	69	31.4	551	1	INVA_PRAVU	ACID BETA-FRUCTOFURANO	6.94e+00
28	69	31.4	638	1	UVRE_HELPY	EXCINUCLEASE ABC SUBUN	6.94e+00
29	69	31.4	3068	1	POLG_PEMVC	GENOME POLYPROTEIN [CO	6.94e+00
30	68	30.9	228	1	PDX3_YEAS	PYRIDOXAMINE 5'-PHOSPH	9.73e+00
31	68	30.9	446	1	MUC_CHICK	IG MU CHAIN C REGION.	9.73e+00
32	68	30.9	486	1	SAHH_TRIVA	ADENOSYLHOMOCYSTEINASE	9.73e+00
33	68	30.9	720	1	ABI3_ARATH	ABSCISIC ACID-INSENSIT	9.73e+00
34	68	30.9	2233	1	COAC_YEAS	ACETYL-COA CARBOXYLASE	9.73e+00
35	68	30.9	3567	1	ERY2_SACER	ERYTHRONOLIDE SYNTHASE	9.73e+00
36	67	30.5	158	1	GLUC_PIG	GLUCAGON PRECURSOR (FR	1.36e+01
37	67	30.5	180	1	GLUC_MOUSE	GLUCAGON PRECURSOR.	1.36e+01
38	67	30.5	180	1	GLUC_OCTDE	GLUCAGON PRECURSOR.	1.36e+01
39	67	30.5	180	1	GLUC_RAT	GLUCAGON PRECURSOR.	1.36e+01
40	67	30.5	180	1	GLUC_BOVIN	GLUCAGON PRECURSOR.	1.36e+01
41	67	30.5	180	1	GLUC_MESAU	GLUCAGON PRECURSOR.	1.36e+01
42	67	30.5	180	1	GLUC_HUMAN	GLUCAGON PRECURSOR.	1.36e+01
43	67	30.5	486	1	YACA_BAGSU	HYPOTHETICAL 55.1 KD P	1.36e+01
44	67	30.5	644	1	KAR9_YEAS	KARYOGAMY PROTEIN KAR9	1.36e+01
45	67	30.5	692	1	YFB3_YEAS	HYPOTHETICAL 78.8 KD P	1.36e+01

ALIGNMENTS

RESULT 1	EXE3_HELHO	STANDARD;	PRT;	39 AA.
ID	P20394;			
AC	01-FEB-1991 (REL. 17, CREATED)			
DT	01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)			
DT	01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)			
DE	EXENDIN-3.			
OS	HELODERMA HORRIDUM (MEXICAN BEADED LIZARD).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;			
OC	SCLEROGLOSSA; ANGILOMORPHA; HELODERMATIDAE; HELODERMA.			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=VENOM;			
RC	MEDLINE; 91056067.			
RA	ENG J., ANDREW P.C., KLEINMAN W.A., SINGH L., RAUFMAN J.-P.;			
RT	"Purification and structure of exendin-3, a new pancreatic			
RT	secretagogue isolated from Heloderma horridum venom.";			
RL	J. BIOL. CHEM. 265:20259-20262(1990).			
CC	-1- FUNCTION: HAS A VIP/SECRETIN-LIKE BIOLOGICAL ACTIVITY. INTERACTS			
CC	WITH THE EXENDIN RECEPTOR.			
CC	-1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.			
DR	PIR; A23674; HMCH32.			
DR	PROSITE; PS00260; GLUCAGON; 1.			
DR	PFAM; PF00123; hormone2; 1.			
DR	HSP; P01274; IGCN.			
KW	GLUCAGON FAMILY; VENOM; AMIDATION.			
FT	MOD_RES 39 39			
SQ	SEQUENCE 39 AA; 4204 MW; AB598FD3 CRC32;			
Query Match	100.0%; Score 220; DB 1; Length 39;			
Best Local Similarity	100.0%; Pred. No. 1.73e-30;			
Matches	31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

Db	9	DLKMEEEAVRLFIWLNKNGPSSGAPPPS 39		
Qy	1	dlsgmeeeavrlfiewlknngpssgappps 31		
RESULT 2	EXE4_HELHU	STANDARD;	PRT;	87 AA.
ID	P26349;			
AC	01-MAY-1992 (REL. 22, CREATED)			
DT	15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)			
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
DE	EXENDIN-4 PRECURSOR.			
OS	HELODERMA SUSPECTUM (GILA MONSTER).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;			

CC SCLEROGLOSSA; ANGIOMORPHA; HELODERMATIDAE; HELODERMA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHEN Y.E., DRUCKER D.J.;
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE OF 48-86.
 RC TISSUE=VENOM;
 RX MEDLINE; 92218391.
 RA ENG J., KLEINMAN W.A., SINGH L., SINGH G., RAUFMAN J.-P.;
 RT "Isolation and characterization of exendin-4, an exendin-3 analogue,
 from Heloderma suspectum venom. Further evidence for an exendin
 receptor on dispersed acini from guinea pig pancreas.";
 RL J. BIOL. CHEM. 267:7402-7405(1992)
 CC -!- FUNCTION: HAS A VIP/SECRETIN-LIKE BIOLOGICAL ACTIVITY. INTERACTS
 WITH THE EXENDIN RECEPTOR.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DB EMBL; U77613; G1916067; -
 DR PIR; A42486; HMGH4G.
 DR PFAM; PF00123; hormone2; 1.
 KW GLUCAGON FAMILY; VENOM; AMIDATION; SIGNAL.
 FT SIGNAL 1 23 POTENTIAL.
 FT PEPTIDE 48 86 EXENDIN-4.
 FT MOD_RES 86 85 AMIDATION (G-87 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 87 AA; 9479 MW; 6C1A8FD5 CRC32;

 Query Match 100.0%; Score 220; DB 1; Length 87;
 Best Local Similarity 100.0%; Pred. No. 1.73e-30;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 DB 56 DLSKQMEEEAVRLFIEWLNKGSSGAPPPS 86
 QY 1 dlskqmeeeavrliwlnkgssgappps 31

 LT 3
 RP LABVS STANDARD; PRT; 2127 AA.
 PL6289;
 DT 01-AUG-1990 (REL. 15, CREATED)
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE RNA POLYMERASE BETA SUBUNIT (EC 2.7.7.48) (LARGE STRUCTURAL PROTEIN)
 DE (L PROTEIN).
 GN L.
 OS RABIES VIRUS (STRAIN SAD B19).
 OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; MONONEGAVIRALES;
 OC RHADDOVIRIDAE; LYSSAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 90223394.
 RA CONZELMANN K.-K., COX J.H., SCHNEIDER L.G., THIEL H.-J.;
 RT "Molecular cloning and complete nucleotide sequence of the attenuated
 rabies virus SAD B19.";
 RL VIROLOGY 175:485-499(1990).
 CC -!- FUNCTION: THIS PROTEIN IS PROBABLY A COMPONENT OF THE ACTIVE
 POLYMERASE. IT MAY FUNCTION IN RNA SYNTHESIS, CAPPING, AS WELL AS
 METHYLATION OF CAPS, AND POLY(A) SYNTHESIS.
 CC -!- SUBUNIT: THOUGHT TO FORM A TRANSCRIPTION COMPLEX WITH THE
 NUCLEOCAPSID (N) PROTEIN.
 CC -!- SIMILARITY: WITH THE L PROTEIN OF OTHER RHADDOVIRUSES AND
 PARAMYXOVIRUSES.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; M31046; G333561; -
 DR PIR; E34746; ZIVNSB.
 KW TRANSFERASE; RNA-DIRECTED RNA POLYMERASE.
 SQ SEQUENCE 2127 AA; 242977 MW; A4044A1E CRC32;

 Query Match 40.0%; Score 88; DB 1; Length 2127;
 Best Local Similarity 42.9%; Pred. No. 6.16e-03;
 Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

 DB 37 NLNSPLIEDPARLMEWLNKG 57
 QY 1 dlskqmeeeavrliwlnkg 21

 RESULT 4
 ID RPPLRABVP STANDARD; PRT; 2142 AA.
 AC P11213;
 DT 01-JUL-1989 (REL. 11, CREATED)
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE RNA POLYMERASE BETA SUBUNIT (EC 2.7.7.48) (LARGE STRUCTURAL PROTEIN)
 DE (L PROTEIN).
 GN L.
 OS RABIES VIRUS (STRAIN PASTEUR / PV).
 OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; MONONEGAVIRALES;
 OC RHADDOVIRIDAE; LYSSAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 88306248.
 RA TORDO N., POCH O., ERMINE A., KEITH G., ROUGEON F.;
 RT "Completion of the rabies virus genome sequence determination: highly
 conserved domains among the L (polymerase) proteins of unsegmented
 negative strand RNA viruses.";
 RL VIROLOGY 165:565-576(1988).
 RN [2]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE; 86233343.
 RA TORDO N., POCH O., ERMINE A., KEITH G., ROUGEON F.;
 RT "Walking along the rabies genome: is the large G-L intergenic region
 a remnant gene?";
 RL PROC. NATL. ACAD. SCI. U.S.A. 83:3914-3918(1986).
 CC -!- FUNCTION: THIS PROTEIN IS PROBABLY A COMPONENT OF THE ACTIVE
 POLYMERASE. IT MAY FUNCTION IN RNA SYNTHESIS, CAPPING, AS WELL AS
 METHYLATION OF CAPS, AND POLY(A) SYNTHESIS.
 CC -!- SUBUNIT: THOUGHT TO FORM A TRANSCRIPTION COMPLEX WITH THE
 NUCLEOCAPSID (N) PROTEIN.
 CC -!- SIMILARITY: WITH THE L PROTEIN OF OTHER RHADDOVIRUSES AND
 PARAMYXOVIRUSES.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; M13215; G333590; -
 DR EMBL; A14671; G492973; -
 DR PIR; A29248; ZIVNVP.
 DR PIR; E24887; E24887.
 KW TRANSFERASE; RNA-DIRECTED RNA POLYMERASE.
 SQ SEQUENCE 2142 AA; 244485 MW; D8D1EB8F CRC32;

 Query Match 40.0%; Score 88; DB 1; Length 2142;
 Best Local Similarity 42.9%; Pred. No. 6.16e-03;

Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Db 37 NLNSPLIEDPARLMLEWLTG 57
 QY 1 dlskqmeeeavrlfiewlkgp 21

RESULT 5
 ID GLUC-ICTPU STANDARD; PRT; 71 AA.
 AC P04093;
 DT 01-NOV-1986 (REL. 03, CREATED)
 DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
 DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
 DE GLUCAGON PRECURSOR (FRAGMENT).
 OS ICTALURUS PUNCTATUS (CHANNEL CATFISH).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 OC TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; SILURIFORMES; ICTALURIDAE;
 OC ICTALURUS.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-PANCREAS;
 MEDLINE; 87156787.
 HOSEIN N.M., MAHREHOLZ A.M., ANDREWS P.C., GURD R.S.;
 "Biological activities of catfish glucagon and glucagon-like
 peptide.";
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 143:87-92(1987).
 RN [2]
 RP SEQUENCE.
 RC TISSUE-PANCREAS;
 MEDLINE; 85157536.
 ANDREWS P.C., RONNER P.;
 RT "Isolation and structures of glucagon and glucagon-like peptide from
 catfish pancreas.";
 RL J. Biol. Chem. 260:3910-3914(1985).
 CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 THE BLOOD SUGAR LEVEL.
 CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -1- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH AMERICAN
 GOOSEFISH SEQUENCES.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PIR; A05166; GCIDC.
 DR PROSITE; PS00260; GLUCAGON; 2.
 DR PFAM; PF00123; hormone2; 2.
 DR HSSP; P01274; 1GCN.
 KW GLUCAGON FAMILY; HORMONE.
 FT NON_TER 1 1
 FT PEPTIDE 1 29 GLUCAGON.
 FT PEPTIDE 38 71 GLUCAGON-LIKE PEPTIDE.
 FT CONFLICT 53 53
 NON_TER 71 71
 SEQUENCE 71 AA; 8173 MW; C49ED93A CRC32;

Query Match 37.7%; Score 83; DB 1; Length 71;
 Best Local Similarity 47.8%; Pred. No. 4.33e-02;
 Matches 11; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Db 46 DVSSYLQEQAKDFIWLKSGQP 68
 QY 1 dlskqmeeeavrlfiewlkgp 23

RESULT 6
 ID GLUC-RANCA STANDARD; PRT; 103 AA.
 AC P15438; P15439; P15440;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE GLUCAGON PRECURSOR (FRAGMENTS).
 OS RANA CATESBEIANA (BULL FROG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
 OC NEOBATRACHIA; RANOIDEA; RANIDAE; RANINAE; RANA.
 RN [1]

SEQUENCE.
 RC TISSUE-PANCREAS;
 MEDLINE; 88257102.
 RA POLLOCK H.G., HAMILTON J.W., ROUSE J.B., EBNER K.E., RAWITCH A.B.;
 RT "Isolation of peptide hormones from the pancreas of the bullfrog
 (Rana catesbeiana). Amino acid sequences of pancreatic polypeptide,
 oxyntomodulin, and two glucagon-like peptides.";
 RL J. BIOL. CHEM. 263:9746-9751(1988).
 CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 THE BLOOD SUGAR LEVEL.
 CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -1- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH OTHER SPECIES
 SEQUENCES.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PIR; B28091; GCGB.
 DR PROSITE; PS00260; GLUCAGON; 3.
 DR PFAM; PF00123; hormone2; 3.
 DR HSSP; P01274; 1GCN.
 KW GLUCAGON FAMILY; HORMONE.
 FT PEPTIDE 1 29 GLUCAGON.
 FT PEPTIDE 1 36 GLUCAGON-36 (OXYNTOMODULIN).
 FT PEPTIDE 39 70 GLUCAGON-LIKE PEPTIDE 1.
 FT NON_CONS 70 71
 FT PEPTIDE 71 103 GLUCAGON-LIKE PEPTIDE 2.
 SQ SEQUENCE 103 AA; 11719 MW; D43EDFC9 CRC32;

Query Match 36.4%; Score 80; DB 1; Length 103;
 Best Local Similarity 43.5%; Pred. No. 1.35e-01;
 Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Db 47 DMSSYLEEKAKEFVDWLKGRP 69
 QY 1 dlskqmeeeavrlfiewlkgp 23

RESULT 7
 ID GLUM-ANGAN STANDARD; PRT; 30 AA.
 AC P41521;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE GLUCAGON-LIKE PEPTIDE (GLP).
 OS ANGUILLA ANGUILLA (EUROPEAN FRESHWATER EEL), AND
 ANGUILLA ROSTRATA (AMERICAN EEL).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 OC TELEOSTEI; ANGUILLIFORMES; ANGUILLIDAE; ANGUILLA.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-PANCREAS;
 MEDLINE; 91340068.
 RA CONLON J.M., ANDREWS P.C., THIM L., MOON T.W.;
 RT "The primary structure of glucagon-like peptide but not insulin has
 been conserved between the American eel, Anguilla rostrata and the
 European eel, Anguilla anguilla.";
 RL GEN. COMP. ENDOCRINOL. 82:23-32(1991).
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PIR; B61125; B61125.
 DR PIR; C61125; C61125.
 DR PROSITE; PS00260; GLUCAGON; 1.
 DR PFAM; PF00123; hormone2; 1.
 DR HSSP; P01274; 1GCN.
 KW GLUCAGON FAMILY; AMIDATION.
 FT MOD_RES 30 30
 SEQUENCE 30 AA; 3376 MW; 27E8C37D CRC32;

Query Match 35.0%; Score 77; DB 1; Length 30;
 Best Local Similarity 38.1%; Pred. No. 4.12e-01;
 Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Db 9 DVSSYLQEQAKDFIWLKSGQP 29
 QY 1 dlskqmeeeavrlfiewlkgp 21

```

RESULT 8
ID GLU2_ORENI STANDARD; PRT; 33 AA.
AC P81027;
DT 01-NOV-1997 (REL. 35, CREATED)
DI 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DL 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DS GLUCAGON II.
CS ORECHROMIS NILOTICUS (NILE TILAPIA) (TILAPIA NILOTICA).
CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
CC TELEOSTEI; EUTELEOSTEI; ACANTHOPTERYGII; PERCOMORPHA; PERCIFORMES;
CC LABROIDEI; CICHLIDAE; TILAPIA.
RN [1]
PP SEQUENCE.
KA MEDLINE; 95384941.
NGUYEN T.M., WRIGHT J.R. JR., NIELSEN P.F., CONLON J.M.;
"Characterization of the pancreatic hormones from the Brockmann body
of the tilapia: implications for islet xenograft studies.";
CC COMP. BIOCHEM. PHYSIOL. 111C:33-44(1995).
CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC THE BLOOD SUGAR LEVEL.
CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC PROSITE; PS00260; GLUCAGON; FALSE_NEG.
DR PFAM; PF00123; hormone2; 1.
DR HSSP; P01274; IGCN.
KW GLUCAGON FAMILY; HORMONE.
SQ SEQUENCE 33 AA; 3731 MW; D0FD0808 CRC32;

Query Match 35.0%; Score 77; DB 1; Length 33;
Best Local Similarity 38.1%; Pred. No. 4.12e-01;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Db 8 DVSSYLQDQAAKEFVSWLTKG 28
|:| :|:|:|:|:|:|:|:|
QY 1 diskmeeeavrlfiewlknq 21

RESULT 9
ID GLUC LEPSP STANDARD; PRT; 78 AA.
AC P095667;
DT 01-MAR-1989 (REL. 10, CREATED)
DI 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DL 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DS GLUCAGON PRECURSOR (FRAGMENT).
CC LEPISTOSTEUS SPATULA (ALLIGATOR GAR) (ATRACTOSTEUS SPATULA).
CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
CC SEMIONOTIFORMES; LEPISTOSTEIDAE; LEPISTOSTEUS.
RN [1]
PP SEQUENCE OF 1-36 AND 45-78.
CC TISSUE-PANCREAS;
RX MEDLINE; 88196798.
RA POLLOCK H.G., KIMMEL J.R., EBNER K.E., HAMILTON J.W., ROUSE J.B.,
RA LANCE V., RAWITCH A.B.;
RT "Isolation of alligator gar (Lepisosteus spatula) glucagon,
RT oxyntomodulin, and glucagon-like peptide: amino acid sequences of
RT oxyntomodulin and glucagon-like peptide.";
RL GEN. COMP. ENDOCRINOL. 69:133-140(1988).
RN [2]
PP PRELIMINARY SEQUENCE OF 1-29.
CC TISSUE-PANCREAS;
RX MEDLINE; 88030594.
RA POLLOCK H.G., KIMMEL J.R., HAMILTON J.W., ROUSE J.B., EBNER K.E.,
RA LANCE V., RAWITCH A.B.;
RT "Isolation and structures of alligator gar (Lepisosteus spatula)
RT insulin and pancreatic polypeptide.";
RL GEN. COMP. ENDOCRINOL. 67:375-382(1987).
CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC THE BLOOD SUGAR LEVEL.
CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.

```

```

CC -!- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH AMERICAN
CC GOSEFISH SEQUENCES.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR PIR; S06339; GCGXA.
DR PROSITE; PS00260; GLUCAGON; 2.
DR PFAM; PF00123; hormone2; 2.
DR HSSP; P01274; IGCN.
KW GLUCAGON FAMILY; HORMONE.
FT PEPTIDE 1 29
FT PEPTIDE 1 36
FT PEPTIDE 45 78
SQ SEQUENCE 78 AA; 8990 MW; 509ED9D3 CRC32;

Query Match 34.5%; Score 76; DB 1; Length 78;
Best Local Similarity 38.1%; Pred. No. 5.93e-01;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Db 53 DVSSYLQDQAAKFFVWLKQG 73
|:| :|:|:|:|:|:|:|:|
QY 1 diskmeeeavrlfiewlknq 21

RESULT 10
ID ACYP_ECOLI STANDARD; PRT; 92 AA.
AC P75877;
DT 15-JUL-1998 (REL. 36, CREATED)
DI 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DL 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DS PUTATIVE ACYLPHOSPHATASE (EC 3.6.1.7) (ACYLPHOSPHATE
DE PHOSPHOHYDROLASE).
GN YCCX.
OS ESCHERICHIA COLI.
CC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
CC ESCHERICHIA.
RN [1]
PP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL SCIENCE 277:1453-1474(1997).
CC -!- CATALYTIC ACTIVITY: AN ACYLPHOSPHATE + H(2)O -> A FATTY ACID ANION
CC + ORTHOPHOSPHATE.
CC -!- SIMILARITY: HIGH, WITH VERTEBRATE ACYLPHOSPHATASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000199; G1787203; -.
CC ECOGENE; EG13726; YCCX.
DR PROSITE; PS00150; ACYLPHOSPHATASE_1; 1.
DR PROSITE; PS00151; ACYLPHOSPHATASE_2; 1.
DR PFAM; PF00708; Acylphosphatase; 1.
DR HSSP; P41500; 2ACY.
KW HYPOTHETICAL PROTEIN; HYDROLASE.
SQ SEQUENCE 92 AA; 10300 MW; 599A4C0C CRC32;

Query Match 34.1%; Score 75; DB 1; Length 92;
Best Local Similarity 50.0%; Pred. No. 8.51e-01;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 51 EEOGVEKLMOWLKSGPRSA 70
|:| :|:|:|:|:|:|:|:|
QY 7 eeeavrlfiewlknqgpssg 26

```

```
RESULT 11
ID GLUC2_CARAU STANDARD; PRT; 121 AA.
AC P79695;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE GLUCAGON PRECURSOR.
OS CARASSIUS AURATUS (GOLDFISH).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; OSTARIOPTERYGII; CYPRINIFORMES; CYPRINOIDEA;
OC CYPRINIDAE; CYPRININAE; CARASSIUS.
RN [1]
RP SEQUENCE FROM N.A.
RA YUEN T.T.H., MOK P.Y., CHOW B.K.C.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC THE BLOOD SUGAR LEVEL.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U06632; G64022; -.
DR EMBL; J00933; G213353; -.
DR PIR; A05150; GCAF2.
DR PROSITE; PS00260; GLUCAGON; 2.
DR PFAM; PF00123; hormone2; 2.
DR HSP; P01274; 1GCM.
KW GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
FT SIGNAL 1 21
FT PEPTIDE 22 49 GRPP (GLICENTINE RELATED POLYPEPTIDE).
FT PEPTIDE 52 80 GLUCAGON II.
FT PEPTIDE 89 119 GLUCAGON-LIKE PEPTIDE II.
SQ SEQUENCE 121 AA; 14171 MW; DFE63061 CRC32;
Query Match 34.1%; Score 75; DB 1; Length 121;
Best Local Similarity 38.1%; Pred. No. 8.51e-01;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
Db 97 DYSSYLQDQAQNFVWLKSGP 117
QY 1 diskqmeeeavrlfiewlkgpp 21
RESULT 12
ID GLUC2_LOPAM STANDARD; PRT; 122 AA.
AC P04092;
DT 01-NOV-1986 (REL. 03, CREATED)
DT 01-NOV-1986 (REL. 03, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE GLUCAGON II PRECURSOR.
OS LOPHIUS AMERICANUS (AMERICAN GOOSEFISH) (ANGLERFISH).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PARACANTHOPTERYGII; LOPHIIFORMES; LOPHIIDAE;
OC LOPHIUS.
RN [1]
RP SEQUENCE FROM N.A.
RA LUND P.K., GOODMAN R.H., MONTMINY M.R., DEE P.C., HABENER J.F.;
RX MEDLINE; 83135785.
RA "Anglerfish islet pre-proglucagon II. Nucleotide and corresponding
RT amino acid sequence of the cDNA."
RL J. BIOL. CHEM. 258:3280-3284(1983).
RN [2]
RP PROCESSING.
RX MEDLINE; 86286913.
RA NOE B.D., ANDREWS P.C.;
RT "Specific glucagon-related peptides isolated from anglerfish islets
RT are metabolic cleavage products of (pre)proglucagon-II."
RL PEPTIDES 7:331-339(1986).
```

```
CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC THE BLOOD SUGAR LEVEL.
CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC -!- IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U06632; G64022; -.
DR EMBL; J00933; G213353; -.
DR PIR; A05150; GCAF2.
DR PROSITE; PS00260; GLUCAGON; 2.
DR PFAM; PF00123; hormone2; 2.
DR HSP; P01274; 1GCM.
KW GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
FT SIGNAL 1 21
FT PEPTIDE 22 49 GRPP (GLICENTINE RELATED POLYPEPTIDE).
FT PEPTIDE 52 80 GLUCAGON II.
FT PEPTIDE 89 119 GLUCAGON-LIKE PEPTIDE II.
SQ SEQUENCE 122 AA; 14171 MW; DFE63061 CRC32;
Query Match 34.1%; Score 75; DB 1; Length 122;
Best Local Similarity 38.1%; Pred. No. 8.51e-01;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
Db 97 DYSSYLQDQAQNFVWLKSGP 117
QY 1 diskqmeeeavrlfiewlkgpp 21
RESULT 13
ID MGMT_HUMAN STANDARD; PRT; 207 AA.
AC P16455;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE METHYLATED-DNA--PROTEIN-CYSTEINE METHYLTRANSFERASE (EC 2.1.1.63) (6-O-
DE METHYLGUANINE-DNA METHYLTRANSFERASE).
GN MGMT.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-8.
RX MEDLINE; 90138892.
RA TANO K., SHIOTA S., COLLIER J., FOOTE R.S., MITRA S.;
RT "Isolation and structural characterization of a cDNA clone encoding
RT the human DNA repair protein for O6-alkylguanine."
RL PROC. NATL. ACAD. SCI. U.S.A. 87:686-690(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA RYDBERG B., SPURR N., KARRAN P.;
RX MEDLINE; 90264461.
RT "cDNA cloning and chromosomal assignment of the human
RT O6-methylguanine-DNA methyltransferase. cDNA expression in
RT Escherichia coli and gene expression in human cells."
RL J. BIOL. CHEM. 265:9563-9569(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90368638.
RA KOIKE G., MAKI H., TAKEYA H., HAYAKAWA H., SEKIGUCHI M.;
RT "Purification, structure, and biochemical properties of human O6-
RT methylguanine-DNA methyltransferase."
RL J. BIOL. CHEM. 265:14754-14762(1990).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90294292.
```



```

RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE: 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL SCIENCE 277:1453-1474 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RA CHUNG E., ALLEN E., ARAUJO R., APARICIO A., DAVIS K., DUNCAN M.,
RA FEDERSPIEL N., HYMAN R., KALMAN S., KOMP C., KURDI O., LEW H., LIN D.,
RA NAMATH A., OEFNER P., ROBERTS D., SCHRAMM S., DAVIS R.W.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000165; G1786818;
DR EMBL: U82598; G1778520;
DR ECOGENE: E013533; YBDN.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 406 AA; 47826 MW; 1DC71FAD CRC32;

```

Query Match 34.1%; Score 75; DB 1; Length 406;
Best Local Similarity 41.2%; Pred. No. 8.51e-01;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

Db 49 ELARQMGKKICVLFIDW 65
   :|||:
QY 1 dlskqmeeeavrifew 17

```

Search completed: Mon Oct 4 15:26:59 1999
Job time : 8 secs.

(™)

Result No.	Query Match	Score	Length	DB ID	Description	Pred. No.
1	50.0	110	265	13	PROGLUCAGON I.	7.43e-06
2	47.7	105	219	13	PROGLUCAGON II.	5.34e-05
3	43.2	95	2127	14	L PROTEIN, RNA DEPENDENCE	2.47e-03
4	41.4	91	379	2	REPHETICAL 42.3 KD P	1.09e-02
5	39.1	86	552	9	REPLICASE.	6.70e-02
6	37.5	81	66	13	PROGLUCAGON (FRAGMENT).	1.53e+00
7	35.0	77	72	13	PROGLUCAGON (FRAGMENT)	1.53e+00
8	35.0	77	72	13	PROGLUCAGON (FRAGMENT)	1.53e+00
9	35.0	77	178	13	GLUCAGON II.	1.53e+00
10	35.0	77	178	13	GLUCAGON I.	1.53e+00
11	34.5	76	300	1	300AA LONG HYPOTHETICA	2.14e+00
12	34.5	76	333	2	INTRINSE/RECOMBINASE.	2.14e+00
13	34.5	76	2185	3	ACETYL COA CARBOXYLASE	2.14e+00
14	34.1	75	664	2	ADENYL-SULPHATE REDU	2.98e+00
15	34.1	75	666	11	PEPTIDYLARGININE DEIMI	2.98e+00
16	34.1	75	665	11	PEPTIDYLARGININE DEIMI	2.98e+00
17	33.6	74	502	5	T25B9.7 PROTEIN.	4.14e+00
18	33.2	73	337	5	C3JE10.8 PROTEIN.	5.74e+00
19	33.2	73	396	14	PUTATIVE VIRAL TEGUMEN	5.74e+00
20	33.2	73	461	5	CODED FOR BY C. ELEGAN	5.74e+00

ALIGNMENTS

RESULT	2				
ID	O42144	PRELIMINARY;	PRT;	219	AA.
AC	O42144;				
DT	01-JAN-1998	(TREMBREL. 05, CREATED)			
DT	01-JAN-1998	(TREMBREL. 05, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998	(TREMBREL. 08, LAST ANNOTATION UPDATE)			
DE	PROGLUCAGON II.				
OS	XENOPUS LAEVIS (AFRICAN CLAWED FROG)..				
OC	EUKARYOTA; METAZOA;				
OC	MESEMBRYOTRACHIA; PIPIDOEA;				
RN	[1]				
RP	SEQUENCE FROM N.A.				

DR EMBL: AF079317; G3378295; -;
 KW HYPOTHETICAL PROTEIN; PLASMIN; ED0127FC CRC32;
 SQ SEQUENCE 379 AA; 42269 MW; 552 AA.
 Query Match 41.4%; Score 91; DB 2; Length 379;
 Best Local Similarity 37.9%; Pred. No. 1.09e-02;
 Matches 11; Conservative 9; Mismatches 8; Indels 1; Gaps 1;
 Db 164 SREMAEMAR-FLEWFAATGGGATPLPG 191
 QY 3 skgmeeeavrlfiewlknngpsgppps 31
 RESULT 5
 ID Q38064 PRELIMINARY; PRT; 552 AA.
 AC Q38064; 552 AA.
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE REPLICASE.
 OS BACTERIOPHAGE PP7.
 OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; LEVIVIRIDAE;
 OC LEVIVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE; 95133199.
 RA OLSTHOORN R.C.L.; GARDE G.; DAYHUFF T.; ATKINS J.F.; VAN DUIN J.;
 RT "Nucleotide sequence of a single-stranded RNA phage from Pseudomonas
 aeruginosa: kinship to coliphages and conservation of regulatory RNA
 structures.";
 RL VIROLOGY 206:611-625(1995).
 DR EMBL; X80191; G517241; -;
 SQ SEQUENCE 552 AA; 63300 MW; 35D63A16 CRC32;
 Query Match 39.1%; Score 86; DB 9; Length 552;
 Best Local Similarity 45.0%; Pred. No. 6.70e-02;
 Matches 9; Conservative 9; Mismatches 0; Indels 2; Gaps 2;
 Db 483 DISKRLDDE-VR-YVDWLRN 500
 QY 1 diskqmeeeavrlfiewlkn 20
 RESULT 6
 ID Q91188 PRELIMINARY; PRT; 66 AA.
 AC Q91188; 66 AA.
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE GLUCAGON (FRAGMENT).
 OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIRDNERI).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;
 OC ONCORHYNCHUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PANCREAS;
 RX MEDLINE; 95295739.
 RA IRWIN D.M.; WONG J.;
 RT "Trout and chicken proglucagon: alternative splicing generates mRNA
 transcripts encoding glucagon-like peptide 2.";
 RL MOL. ENDOCRINOL. 9:267-277(1995).
 DR EMBL; U19913; G736361; -;
 DR PFAM; PF00123; hormone2; 2.
 FT NONTER 1
 SQ SEQUENCE 66 AA; 7680 MW; 62C576E2 CRC32;
 Query Match 35.0%; Score 77; DB 13; Length 66;
 Best Local Similarity 38.1%; Pred. No. 1.53e+00;
 Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
 Db 41 DVSTYLQDQAQKDFVSWLKS 61
 QY 1 dvstylqddqaaqkdfvswlks 61
 RESULT 7
 ID Q85863 PRELIMINARY; PRT; 379 AA.
 AC Q85863; 379 AA.
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 42.3 KD PROTEIN.
 OS SPHINGOMONAS AROMATICIVORANS.
 OC PLASMAID PHLA.
 OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; ZYMOMONAS GROUP;
 OC SPHINGOMONAS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R199;
 RA KIMINE M.F.; STILLWELL L.C.; WONG K.-K.; THURSTON S.J.; SISK E.C.;
 PA SENSEN C.W.; GAASTERLAND T.; SAFFER J.D.; FREDRICKSON J.K.;
 RT "Complete sequence of a 184 kb catabolic plasmid from Spingomonas
 aromaticivorans strain F199.";
 RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: AF079317; G3378295; -;
 KW HYPOTHETICAL PROTEIN; PLASMIN; ED0127FC CRC32;
 SQ SEQUENCE 379 AA; 42269 MW; 552 AA.
 Query Match 41.4%; Score 91; DB 2; Length 379;
 Best Local Similarity 37.9%; Pred. No. 1.09e-02;
 Matches 11; Conservative 9; Mismatches 8; Indels 1; Gaps 1;
 Db 164 SREMAEMAR-FLEWFAATGGGATPLPG 191
 QY 3 skgmeeeavrlfiewlknngpsgppps 31
 RESULT 5
 ID Q38064 PRELIMINARY; PRT; 552 AA.
 AC Q38064; 552 AA.
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE REPLICASE.
 OS BACTERIOPHAGE PP7.
 OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; LEVIVIRIDAE;
 OC LEVIVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE; 95133199.
 RA OLSTHOORN R.C.L.; GARDE G.; DAYHUFF T.; ATKINS J.F.; VAN DUIN J.;
 RT "Nucleotide sequence of a single-stranded RNA phage from Pseudomonas
 aeruginosa: kinship to coliphages and conservation of regulatory RNA
 structures.";
 RL VIROLOGY 206:611-625(1995).
 DR EMBL; X80191; G517241; -;
 SQ SEQUENCE 552 AA; 63300 MW; 35D63A16 CRC32;
 Query Match 39.1%; Score 86; DB 9; Length 552;
 Best Local Similarity 45.0%; Pred. No. 6.70e-02;
 Matches 9; Conservative 9; Mismatches 0; Indels 2; Gaps 2;
 Db 483 DISKRLDDE-VR-YVDWLRN 500
 QY 1 diskqmeeeavrlfiewlkn 20
 RESULT 6
 ID Q91188 PRELIMINARY; PRT; 66 AA.
 AC Q91188; 66 AA.
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE GLUCAGON (FRAGMENT).
 OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIRDNERI).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;
 OC ONCORHYNCHUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PANCREAS;
 RX MEDLINE; 95295739.
 RA IRWIN D.M.; WONG J.;
 RT "Trout and chicken proglucagon: alternative splicing generates mRNA
 transcripts encoding glucagon-like peptide 2.";
 RL MOL. ENDOCRINOL. 9:267-277(1995).
 DR EMBL; U19913; G736361; -;
 DR PFAM; PF00123; hormone2; 2.
 FT NONTER 1
 SQ SEQUENCE 66 AA; 7680 MW; 62C576E2 CRC32;
 Query Match 35.0%; Score 77; DB 13; Length 66;
 Best Local Similarity 38.1%; Pred. No. 1.53e+00;
 Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
 Db 41 DVSTYLQDQAQKDFVSWLKS 61
 QY 1 dvstylqddqaaqkdfvswlks 61
 RESULT 7
 ID Q85863 PRELIMINARY; PRT; 379 AA.
 AC Q85863; 379 AA.
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 42.3 KD PROTEIN.
 OS SPHINGOMONAS AROMATICIVORANS.
 OC PLASMAID PHLA.
 OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; ZYMOMONAS GROUP;
 OC SPHINGOMONAS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R199;
 RA KIMINE M.F.; STILLWELL L.C.; WONG K.-K.; THURSTON S.J.; SISK E.C.;
 PA SENSEN C.W.; GAASTERLAND T.; SAFFER J.D.; FREDRICKSON J.K.;
 RT "Complete sequence of a 184 kb catabolic plasmid from Spingomonas
 aromaticivorans strain F199.";
 RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.

AC Q91189; Q92168;
DT 01-NOV-1996 (TREMREL. 01, CREATED)

1

AC Q91189; Q92168;
DT 01-NOV-1996 (TREMREL. 01, CREATED)

```

OS PYROCOCUS HORIKOSHII.
OC ARCHAEA; EUKYARCHAEOTA; THERMOCOCCALES; THERMOCOCCACEAE; PYROCOCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3.
RX MEDLINE; 98344137.
RA KAWARABAYASHI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,
KA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOVAMA A., NAGAI Y.,
KA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSUDA N., OGUCHI A.,
KA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,
KA KIKUCHI H.;
*Complete Sequence and Gene Organization of the Genome of a
RT Hyper-thermophilic Archaeobacterium, Pyrococcus horikoshii OT3.*;
RL DNA RES. 5:55-76(1998).
DR EMBL; AP000006; D1031532; -.
PROSITE; PS00782; TFIIB; 2.
INITIATION FACTOR.
SQ SEQUENCE 300 AA; 34097 MW; 6E17BB64 CRC32;

Query Match 34.5%; Score 76; DB 1; Length 300;
Best Local Similarity 50.0%; Pred. No. 2.14e+00;
Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Db 125 LPKHVEEERARLYREAVRKG 144
QY 2 lskqmeeeavrlfiewlknq 21

RESULT 12
ID Q48338 PRELIMINARY; PRT; 333 AA.
AC Q48338;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE INTEGRASE/RECOMBINASE.
OS LACTOBACILLUS DELBRUECKII.
OG PLASMID PWS58.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; LACTOBACILLACEAE;
OC LACTOBACILLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WS58;
RA KLEIN J.R., HENRICH B.;
SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
EMBL; Z50864; G971481; -.
PFAM; PF00589; Phage_integrase; 1.
PLASMID.
SQ SEQUENCE 333 AA; 38147 MW; 6704796A CRC32;

Query Match 34.5%; Score 76; DB 2; Length 333;
Best Local Similarity 30.4%; Pred. No. 2.14e+00;
Matches 7; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Db 73 QTSESTIKLYMOMLENGROPST 95
QY 5 qmeeeavrlfiewlknqgppssa 27

RESULT 13
ID Q12721 PRELIMINARY; PRT; 2185 AA.
AC Q12721;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ACETYL COA CARBOXYLASE (EC 6.4.1.2) (ACETYL-COA CARBOXYLASE).
GN ACC.
OS USTILAGO MAYDIS (SMUT FUNGUS).
OC EUKARYOTA; FUNGI; BASIDIOMYCOTA; USTILAGINOMYCETES;
OC USTILAGINOMYCETIDAE; USTILAGINALES; USTILAGINACEAE; USTILAGO.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TM103761;

```

```

RX MEDLINE; 95086936.
RA BAILEY A.M., KEON J.P.R., OWEN J., HARGREAVES J.A.;
RT "The ACCI gene, encoding acetyl-CoA carboxylase, is essential for
growth in Ustilago maydis.";
RL MOL. GEN. GENET. 249:191-201(1995).
CC -|- CATALYTIC ACTIVITY: ATP + ACETYL-COA + HCO(3)(-) = ADP +
PHOSPHATE + MALONYL-COA.
CC -|- COFACTOR: BIOTIN
CC EMBL; 446886; G600098; -.
DR PFAM; PF00289; CFSase_L_chain; 1.
DR PFAM; PF00364; biotin_req_enzy; 1.
DR PFAM; PF01039; Carboxyl_trans; 1.
KW LIGASE.
SQ SEQUENCE 2185 AA; 240029 MW; 84AA60F1 CRC32;

Query Match 34.5%; Score 76; DB 3; Length 2185;
Best Local Similarity 34.5%; Pred. No. 2.14e+00;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Db 1763 LTAODDLDAVRSFVNWISVPAQGGGLP 1791
QY 2 lskqmeeeavrlfiewlknqgppsgapp 30

RESULT 14
ID Q59339 PRELIMINARY; PRT; 664 AA.
AC Q59339;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE ADENYL-SULPHATE REDUCTASE ALFA-SUBUNIT (EC 1.8.9.2)
DE (ADENYL-SULPHATE REDUCTASE).
GN APSA.
OS DESULFOVIBRIO VULGARIS.
OC BACTERIA; PROTEOBACTERIA; DELTA SUBDIVISION; DESULFOVIBRIO.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HILDENBOROUGH NCIMB 8303;
RA VAN DEN BERG W., VAN DONGEN W., HAGEN W.;
SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HILDENBOROUGH NCIMB 8303;
RX MEDLINE; 94362893.
RA SPEICH N., DAHL C., HEISIG P., KLEIN A., LOTTSPEICH F., STETTER K.,
RA TRUPER H.;
RT "Adenylsulphate reductase from the sulphate-reducing archaeon
Archaeoglobus fulgidus: cloning and characterization of the genes and
comparison of the enzyme with other iron-sulphur flavoproteins.";
RL MICROBIOLOGY 140:1273-1284(1994).
CC -|- CATALYTIC ACTIVITY: AMP + SULFITE + ACCEPTOR = ADENYL-SULFATE +
REDUCED ACCEPTOR.
CC -|- COFACTOR: FAD; IRON.
CC EMBL; 569372; E221398; -.
DR PFAM; PF00890; FAD_binding_2; 1.
KW OXIDOREDUCTASE.
SQ SEQUENCE 664 AA; 74627 MW; ED9E9E3E CRC32;

Query Match 34.1%; Score 75; DB 2; Length 664;
Best Local Similarity 41.2%; Pred. No. 2.98e+00;
Matches 7; Conservative 8; Mismatches 1; Indels 1; Gaps 1;

Db 108 DLGRHVDD-SVHLFEW 123
QY 1 lskqmeeeavrlfiew 17

RESULT 15
ID O88807 PRELIMINARY; PRT; 666 AA.
AC O88807;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

```

DE PEPTIDYLARGININE DEIMINASE TYPE IV (EC 3.5.3.15)
 DE (PROTEIN-ARGININE DEIMINASE).
 GN PAD-R4.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
 OC SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 98409324.
 RA ISHIGAMI A., KURAMOTO M., YAMADA M., WATANABE K., SENSU T.;
 RT "Molecular cloning of two novel types of peptidylarginine deiminase
 RT cDNAs from retinoic acid-treated culture of a newborn rat
 RT keratinocyte cell line".
 RL FEBS LETT. 433:113-118(1998).
 CC -!- CATALYTIC ACTIVITY: PROTEIN L-ARGININE + H(2)O -> PROTEIN
 CC L-CITRULLINE + NH(3).
 DR EMBL; AB010999; D1033062; -.
 KW HYDROLASE.
 SQ SEQUENCE 566 AA; 74467 MW; 3E64EDDD CRC32;

Query Match 34.1%; Score 75; DB 11; Length 666;
 est Local Similarity 47.1%; Pred. No. 2.98e+00;
 tches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 451 LGQQVQAPVRLFSDWL 467
 QY 1:1:|||||:
 2 lskqmeeeavrlfiewl 18

Search completed: Mon Oct 4 15:27:33 1999
 Job time: 17 secs.

MOHAM-312-CLAIM82B.PEP

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MOHAM-312-CLAIM82B.PEP protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Oct 4 15:26:14 1999; MasPar time 11.75 Seconds
bular output not generated. 56.092 Million cell updates/sec

File: >MOHAM-312-CLAIM82B.PEP
Description: (1-31) from moham312177.ppe
Perfect Score: 231
Sequence: 1 hgegtfslskmqeeavrlfiewlknngy 31

Scoring table: PAM 150
Gap 11

Searched: 170751 seqs, 2126608 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-genseq35

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Statistics: Mean 24.523; Variance 105.941; scale 0.231

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	231	100.0	31	14	Heloderma suspectum e	1.93e-12
2	219	94.8	30	39	Exendin-4 (1-30) for	2.44e-11
3	219	94.8	31	14	Heloderma suspectum e	2.44e-11
4	219	94.8	39	14	Heloderma suspectum e	2.44e-11
5	219	94.8	39	39	Exendin-4, for use in	2.44e-11
6	219	94.8	39	30	Gila monster exendin-	2.44e-11
7	219	94.8	87	35	Heloderma suspectum p	2.44e-11
8	212	91.8	30	29	H. horridum exendin-4	1.06e-10
9	212	91.8	39	39	Exendin-3, for use in	1.06e-10
10	212	91.8	39	30	Gila monster exendin-	1.06e-10
11	212	91.8	39	14	Heloderma horridum ex	1.06e-10
12	207	89.6	30	29	H. horridum exendin-4	3.03e-10
13	205	88.7	28	39	W39309	4.60e-10
14	205	88.7	30	29	H. horridum exendin-3	4.60e-10
15	205	88.7	30	29	W39301	4.60e-10
16	203	87.9	30	29	H. horridum exendin-4	6.99e-10

17	202	87.4	30	29	W39304	H. horridum exendin-4	8.62e-10
18	201	87.0	30	29	W39303	H. horridum exendin-4	1.06e-09
19	201	87.0	30	29	W39308	H. horridum exendin-4	1.06e-09
20	201	87.0	30	29	W39306	H. horridum exendin-4	1.06e-09
21	200	86.6	30	29	W39383	H. horridum exendin-3	1.31e-09
22	199	86.1	30	29	W39367	H. horridum exendin-4	1.61e-09
23	199	86.1	30	29	W39311	H. horridum exendin-4	1.61e-09
24	198	85.7	30	29	W39349	H. horridum exendin-4	1.98e-09
25	197	85.3	30	29	W39358	H. horridum exendin-4	2.44e-09
26	197	85.3	30	29	W39351	H. horridum exendin-4	2.44e-09
27	197	85.3	30	29	W39347	H. horridum exendin-4	2.44e-09
28	197	85.3	30	29	W39361	H. horridum exendin-4	2.44e-09
29	196	84.8	30	29	W39345	H. horridum exendin-4	3.01e-09
30	196	84.8	30	29	W39341	H. horridum exendin-4	3.01e-09
31	196	84.8	30	29	W39310	H. horridum exendin-4	3.01e-09
32	195	84.4	30	29	W39317	H. horridum exendin-4	3.71e-09
33	195	84.4	30	29	W39369	H. horridum exendin-3	3.71e-09
34	195	84.4	30	29	W39343	H. horridum exendin-4	3.71e-09
35	195	84.4	30	29	W39331	H. horridum exendin-4	3.71e-09
36	194	84.0	30	29	W39370	H. horridum exendin-3	4.56e-09
37	194	84.0	30	29	W39327	H. horridum exendin-4	4.56e-09
38	194	84.0	30	29	W39420	H. horridum exendin-3	4.56e-09
39	194	84.0	30	29	W39378	H. horridum exendin-3	4.56e-09
40	194	84.0	30	29	W39319	H. horridum exendin-4	4.56e-09
41	194	84.0	30	29	W39305	H. horridum exendin-3	4.56e-09
42	194	84.0	39	39	W61773	Leu(14), Phe(25)-exen	4.56e-09
43	193	83.5	30	29	W39332	H. horridum exendin-4	5.62e-09
44	193	83.5	30	29	W39354	H. horridum exendin-4	5.62e-09
45	193	83.5	30	29	W39380	H. horridum exendin-3	5.62e-09

ALIGNMENTS

RESULT 1
ID R80544 standard; peptide; 31 AA.
AC R80544;
DT 27-FEB-1996 (first entry)
DE Heloderma suspectum exendin-4 residues 1-31-Tyr31.
KW Exendin-4; residues 1-31; Y-31-Exendin-4(1-31); diabetes mellitus;
KW hyperglycaemia; Tyr31; insulinotropic peptide.
OS Heloderma suspectum.
PN US5424286-A.
PD 13-JUN-1995.
PF 24-MAY-1993; 066480.
PR 24-MAY-1993; US-066480.
PA (ENGJ/) ENG J.
PI Eng J;
DR WPI; 95-262627/34
PT Stimulating/inhibiting insulin release with exendin polypeptide(s) -
PT for treating diabetes mellitus and preventing hyperglycaemia.
PS Claim 2; Columns 13-14; 17pp; English.
CC R80544 is the Heloderma suspectum exendin-4 residues 1-31, where
CC the native Pro31 has been replaced with a Tyr residue. It is an
CC insulinotropic peptide, and can therefore be used in the treatment of
CC diabetes mellitus (types I or II), and for the prevention of
CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
CC and insulin-(in)dependent mechanisms.
SQ Sequence 31 AA;

Query Match 100.0%; Score 231; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.93e-12;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtfslskmqeeavrlfiewlknngy 31
|||||
QY 1 hgegtfslskmqeeavrlfiewlknngy 31

RESULT 2
ID W61771 standard; peptide; 30 AA.
AC W61771;
DT 29-MAR-1999 (first entry)
DE Exendin-4 (1-30) for use in treating disorders related to food intake.

KW Exendin; obesity; type II diabetes; eating disorders; cardiac disease;
 OS insulin resistance syndrome; elevated plasma glucose level; agonist.
 Heloderma suspectum.
 FH Key Location/Qualifiers
 FT Modified_site 30
 /note= "optionally the C-terminal is in amide form"
 WO9830231-A1.
 16-JUL-1998.
 PD 07-JAN-1998; U00449.
 PR 14-NOV-1997; US-066029.
 PR 07-JAN-1997; US-034905.
 PR 08-AUG-1997; US-055404.
 PR 14-NOV-1997; US-065442.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Beeley NRA, Bhavsar S, Prickett KS;
 DR WPI; 98-398796/34.
 Reducing food intake by administering exendins or their
 analogues - for treatment of e.g. obesity, type II diabetes,
 eating disorders and insulin resistance
 Claims 18, 26; Page 11; 214pp; English.
 PS The invention relates to a new method for treating disorders that
 CC are alleviated by reducing food intake, in particular obesity, type
 CC II diabetes, eating disorders, insulin resistance syndrome, elevated
 CC plasma glucose levels, or the risk of cardiac disease. The method
 CC comprises administering an exendin or an exendin agonist. The method
 CC reduces appetite and lowers plasma lipid levels. It inhibits food
 CC consumption as effectively as amylin or cholecystokinin but has a much
 CC longer-lasting action (still effective after 6 hours in a mouse model).
 CC The present sequence is that of exendin-4 (1-30) or its amide which is
 CC one of the preferred compounds for use in the method.
 SQ Sequence 30 AA;

Query Match 94.8%; Score 219; DB 39; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.44e-11;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtfslskqmeeeavrlfiewlknng 30
 |||||
 QY 1 hgegtfslskqmeeeavrlfiewlknng 30

RESULT 3

ID R80543 standard; peptide; 31 AA.
 AC R80543;
 DE Heloderma suspectum exendin-4 residues 1-31 (Exendin-4(1-31)).
 DT Exendin-4; residues 1-31; Exendin-4(1-31); diabetes mellitus;
 KW hyperglycaemia; insulinotropic peptide.
 OS Heloderma suspectum.
 PN US5424286-A.
 PD 13-JUN-1995.
 PF 24-MAY-1993; 066480.
 PR 24-MAY-1993; US-066480.
 PA (ENGJ/) ENG J.
 PI Eng J;
 DR WPI; 95-262627/34.
 PT Stimulating/inhibiting insulin release with exendin polypeptide(s) -
 PT for treating diabetes mellitus and preventing hyperglycaemia.
 PS Claim 1; Columns 13-14; 17pp; English.
 CC R80543 is the Heloderma suspectum exendin-4 residues 1-31. It is an
 CC insulinotropic peptide, and can therefore be used in the treatment of
 CC diabetes mellitus (types I or II), and for the prevention of
 CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
 CC and insulin-(in)dependent mechanisms.
 SQ Sequence 31 AA;

Query Match 94.8%; Score 219; DB 14; Length 31;
 Best Local Similarity 100.0%; Pred. No. 2.44e-11;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtfslskqmeeeavrlfiewlknng 30
 |||||
 QY 1 hgegtfslskqmeeeavrlfiewlknng 30

RESULT 4
 ID R80546 standard; peptide; 39 AA.
 AC R80546;
 DT 27-FEB-1996 (first entry)
 DE Heloderma suspectum exendin-4.
 KW Exendin-4; diabetes mellitus; hyperglycaemia; insulinotropic peptide.
 OS Heloderma suspectum.
 PN US5424286-A.
 PD 13-JUN-1995.
 PF 24-MAY-1993; 066480.
 PR 24-MAY-1993; US-066480.
 PA (ENGJ/) ENG J.
 PI Eng J;
 DR WPI; 95-262627/34.
 PT Stimulating/inhibiting insulin release with exendin polypeptide(s) -
 PT for treating diabetes mellitus and preventing hyperglycaemia.
 PS Claim 6; Columns 13-14; 17pp; English.
 CC R80546 is Heloderma suspectum exendin-4. It is an
 CC insulinotropic peptide, and can therefore be used in the treatment of
 CC diabetes mellitus (types I or II), and for the prevention of
 CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
 CC and insulin-(in)dependent mechanisms.
 SQ Sequence 39 AA;

Query Match 94.8%; Score 219; DB 14; Length 39;
 Best Local Similarity 100.0%; Pred. No. 2.44e-11;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtfslskqmeeeavrlfiewlknng 30
 |||||
 QY 1 hgegtfslskqmeeeavrlfiewlknng 30

RESULT 5

ID W61770 standard; peptide; 39 AA.
 AC W61770;
 DT 29-MAR-1999 (first entry)
 DE Exendin-4, for use in treating disorders related to food intake.
 KW Exendin; obesity; type II diabetes; eating disorders; cardiac disease;
 OS Heloderma suspectum.
 PN WO9830231-A1.
 PD 16-JUL-1998.
 PF 07-JAN-1998; U00449.
 PR 14-NOV-1997; US-066029.
 PR 07-JAN-1997; US-034905.
 PR 08-AUG-1997; US-055404.
 PR 14-NOV-1997; US-065442.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Beeley NRA, Bhavsar S, Prickett KS;
 DR WPI; 98-398796/34.
 PT Reducing food intake by administering exendins or their
 PT analogues - for treatment of e.g. obesity, type II diabetes,
 PT eating disorders and insulin resistance
 PS Claims 17, 25; Page 8; 214pp; English.
 CC The invention relates to a new method for treating disorders that
 CC are alleviated by reducing food intake, in particular obesity, type
 CC II diabetes, eating disorders, insulin resistance syndrome, elevated
 CC plasma glucose levels, or the risk of cardiac disease. The method
 CC comprises administering an exendin or an exendin agonist. The method
 CC reduces appetite and lowers plasma lipid levels. It inhibits food
 CC consumption as effectively as amylin or cholecystokinin but has a much
 CC longer-lasting action (still effective after 6 hours in a mouse model).
 CC The present sequence is that of exendin-4 which is one of the preferred
 CC compounds for use in the method.
 SQ Sequence 39 AA;

Query Match 94.8%; Score 219; DB 39; Length 39;
 Best Local Similarity 100.0%; Pred. No. 2.44e-11;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtfslskqmeeeavrlfiewlknng 30
 QY 1 hgegtfslskqmeeeavrlfiewlknng 30

RESULT 6

ID W47609 standard; peptide; 39 AA.
 AC W47609.
 DT 03-JUL-1998 (first entry)
 DE Gila monster extendin-4.
 KW Extendin agonist; gastric motility; gastric emptying; treatment;
 KW spasm; postprandial dumping syndrome; postprandial hyperglycaemia;
 KW type 1 diabetes; impaired glucose tolerance; toxin ingestion;
 KW obesity; Gila monster venom; extendin-4.
 OS Heloderma suspectum.
 FH Key Location/Qualifiers
 FT Modified_site 39 /note- "amidated"
 FT W09805351-A1.
 PN 12-FEB-1998.
 PD 08-AUG-1997; UL4199.
 PF 08-AUG-1996; US-694954.
 PR (AMYL-) AMYLIN PHARM INC.
 PA Beelley NRA, Gedulin B, Prickett KS, Young AA;
 PI WPI; 98-145351/13.
 PT Regulating gastrointestinal motility using extendins or their
 PT agonists - for treating spasm, diabetic postprandial hyperglycaemia,
 PT impaired glucose tolerance etc., also in diagnostic investigations
 PT Claims 20 and 21; Fig 1; 70pp; English.
 PS W47549 describes a generic extendin agonist, provided that it does
 CC have the formula of either extendin-3 (W47608) or extendin-4
 CC (W47609).

CC Extendin agonists, which reduce gastric motility and delay gastric
 CC emptying, can be used to treat spasm (where associated with acute
 CC diverticulitis or disorders of the biliary tract or sphincter of
 CC Oddi), postprandial dumping syndrome and hyperglycaemia
 CC (particularly associated with type 2 diabetes), type 1 diabetes,
 CC impaired glucose tolerance, toxin ingestion (an extendin agonist is
 CC administered to prevent stomach contents passing into the
 CC intestines, then the stomach pumped) and obesity. They can also be
 CC administered to subjects undergoing gastrointestinal diagnostic
 CC investigation, particularly radiological or by magnetic resonance
 CC imaging.
 CC Extendins, components of Gila monster venom, have some sequence
 CC similarity to glucagon-like peptides (GLP). They are GLP agonists
 CC and have been suggested (US5424286) for treatment of diabetes and
 CC prevention of hyperglycaemia.
 SQ Sequence 39 AA;

Query Match 94.8%; Score 219; DB 30; Length 39;
 Best Local Similarity 100.0%; Pred. No. 2.44e-11;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtfslskqmeeeavrlfiewlknng 30
 QY 1 hgegtfslskqmeeeavrlfiewlknng 30

RESULT 7

ID W70288 standard; Protein; 87 AA.
 AC W70288;
 DT 06-NOV-1998 (first entry)
 DE Heloderma suspectum proextendin peptide.
 KW Heloderma suspectum proextendin; extendin N-terminal peptide; ENTp;
 KW extendin 4 peptide; extendin 3 gene; Heloderma horridum; metabolic disease;
 KW drug screening; endocrine tumour; organ failure; cell metabolism;
 KW diabetes; reptilian venom peptide.
 OS Heloderma suspectum.
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT /note- "Signal peptide"
 FT Peptide 1..47
 FT /note- "ENTp"

FT Peptide 48..87
 FT Cleavage_site 46..47 /note- "Extendin 4"
 FT /note- "Dipeptidyl peptidase cleavage site"

PN W09835033-A1.
 PD 13-AUG-1998.
 PF 04-FEB-1998; CA0071.
 PR 07-FEB-1997; GB-002582.
 PR 05-FEB-1997; US-037412.
 PA (ONEO-) 1149336 ONTARIO INC.
 PI Drucker DJ;

PI WPI; 98-447230/38.
 DR N-PSDB; V33163.
 PT New nucleic acid encoding proextendin - used to diagnose and treat,
 PT e.g. endocrine tumours, also to treat poisoning by reptile venom
 PT Claim 3; Fig 2; 26pp; English.
 PS The Heloderma suspectum proextendin peptide is encoded by its cDNA
 CC which was isolated from a H. suspectum salivary gland cDNA library.
 CC The proextendin protein comprises of a novel extendin N-terminal
 CC peptide (ENTP) linked to the N-terminus of the extendin 4 peptide
 CC by a consensus dipeptidyl peptidase cleavage site. The proextendin
 CC cDNA can be used to clone or identify related sequences (e.g. the
 CC extendin 3 gene of Heloderma horridum, mutant alleles and proextendin
 CC gene regulatory defects associated with metabolic disease) and species
 CC homologues (e.g. for developing animal models for drug screening).
 CC The proextendin peptide can be used to raise antibodies. Anti-proextendin
 CC antibodies are claimed to be useful for diagnosing conditions associated
 CC with altered levels of proextendin (e.g. endocrine tumours and organ
 CC failure), for identifying other regulators of cell metabolism, in drug
 CC screens and for treating metabolic diseases (e.g. diabetes) and for
 CC neutralising, or detecting, reptilian venom peptides.
 SQ Sequence 87 AA;

Query Match 94.8%; Score 219; DB 35; Length 87;
 Best Local Similarity 100.0%; Pred. No. 2.44e-11;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 48 hgegtfslskqmeeeavrlfiewlknng 77
 QY 1 hgegtfslskqmeeeavrlfiewlknng 30

RESULT 8

ID W39302 standard; peptide; 30 AA.
 AC W39302;
 DT 05-JUN-1998 (first entry)
 DE H. horridum extendin-4 peptide.
 KW Extendin-3; extendin 4; diabetes; insulin; secretion; biosynthesis;
 KW glucagon reduction; hypoglycaemia; glucose; treatment.
 OS Heloderma horridum.
 FH Key Location/Qualifiers
 FT Modified_site 30
 FT /note- "This residue can be any amino acid except Gly"

FT W09746584-A1.
 PN 11-DEC-1997.
 PD 05-JUN-1997; E02930.
 PF 13-SEP-1996; DE-037230.
 PR 05-JUN-1996; DE-022502.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Goeke B, Goeke R, Hoffmann B;
 PI WPI; 98-042119/04.

DR Truncated versions of extendin peptide(s) for treating diabetes -
 FT increase secretion and biosynthesis of insulin, but reduce those of
 FT glucagon, and do not induce hypoglycaemia
 PS Claim 1; Page 4; 150pp; English.
 CC This peptide is a fragment of extendin-4 isolated from Heloderma
 CC horridum. This peptide and its salts, esters and derivatives can be
 CC used to treat diabetes mellitus. They stimulate biosynthesis and
 CC secretion of insulin, but have the opposite effect on glucagon, and
 CC independent of this activity can increase peripheral glucose utilisation.
 CC Extendin-3 and extendin-4 are only active when blood sugar levels are
 CC high, so they will not induce hypoglycaemia. Compared with glucagon-like

CC peptide 1 (GLP1) and the known extendins, they are more active (effective
 CC at lower doses), more stable to degradation and metabolism and have a
 CC longer lasting effect. Truncated forms of this peptide can be made more
 CC economically than full length versions.
 SQ Sequence 30 AA;

Query Match 91.8%; Score 212; DB 29; Length 30;
 Best Local Similarity 100.0%; Pred. No. 1.06e-10;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hsgdftsdlskqmeeeavrlfiewlknng 29
 QY 1 hgegtftsdlskqmeeeavrlfiewlknng 29

RESULT 9
 W61769 standard; peptide; 39 AA.

DE 29-MAR-1999 (first entry)
 KW Extendin-3, for use in treating disorders related to food intake.
 KW Insulin resistance syndrome; eating disorders; cardiac disease;
 OS Heloderma horridum.
 PN W09830231-A1.
 PD 16-JUL-1998.
 PF 07-JAN-1998; U00449.
 PR 14-NOV-1997; US-066029.
 PR 07-JAN-1997; US-034905.
 PR 08-AUG-1997; US-055404.
 PR 14-NOV-1997; US-065442.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Beeley NRA, Bhavsar S, Prickett KS;
 DR WPI; 98-398796/34.
 PT Reducing food intake by administering extendins or their
 PT analogues - for treatment of e.g. obesity, type II diabetes,
 PT eating disorders and insulin resistance
 PS Claims 15, 24; Page 8; 214pp; English.
 CC The invention relates to a new method for treating disorders that
 CC are alleviated by reducing food intake, in particular obesity, type
 CC II diabetes, eating disorders, insulin resistance syndrome, elevated
 CC plasma glucose levels, or the risk of cardiac disease. The method
 CC comprises administering an extendin or an extendin agonist. The treatment
 CC reduces appetite and lowers plasma lipid levels. It inhibits food
 CC consumption as effectively as amylin or cholecystokinin but has a much
 CC longer-lasting action (still effective after 6 hours in a mouse model).
 CC The present sequence is that of extendin-3 which is one of the preferred
 CC compounds for use in the method.
 SQ Sequence 39 AA;

Query Match 91.8%; Score 212; DB 39; Length 39;
 Best Local Similarity 93.3%; Pred. No. 1.06e-10;
 Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 hsgdftsdlskqmeeeavrlfiewlknng 30
 QY 1 hgegtftsdlskqmeeeavrlfiewlknng 30

RESULT 10
 W47608 standard; peptide; 39 AA.

DE 01-JUL-1998 (first entry)
 DE Gila monster extendin-3.
 KW Extendin agonist; gastric motility; gastric emptying; treatment;
 KW spasm; postprandial dumping syndrome; postprandial hyperglycaemia;
 KW type I diabetes; impaired glucose tolerance; toxin ingestion;
 KW obesity; Gila monster venom; extendin-3.
 CS Heloderma horridum.
 CS WPI; 98-398796/34.

FT Modified_site 39 Location/Qualifiers
 FT /note= "amidated"
 PN W09805351-A1.
 PD 12-FEB-1998.

PF 08-AUG-1997; U14199.
 PR 08-AUG-1996; US-694954.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Beeley NRA, Gedulin B, Prickett KS, Young AA;
 DR WPI; 98-145351/13.
 PT Regulating gastrointestinal motility using extendins or their
 PT agonists - for treating spasm, diabetic postprandial hyperglycaemia,
 PT impaired glucose tolerance etc., also in diagnostic investigations
 PS Claims 20 and 21; Fig 1; 70pp; English.
 CC W47549 describes a generic extendin agonist, provided that it does
 CC have the formula of either extendin-3 (W47608) or extendin-4
 CC (W47609).
 CC Extendin agonists, which reduce gastric motility and delay gastric
 CC emptying, can be used to treat spasm (where associated with acute
 CC diverticulitis or disorders of the biliary tract or sphincter of
 CC Oddi), postprandial dumping syndrome and hyperglycaemia
 CC (particularly associated with type 2 diabetes), type 1 diabetes,
 CC impaired glucose tolerance, toxin ingestion (an extendin agonist is
 CC administered to prevent stomach contents passing into the
 CC intestines, then the stomach pumped) and obesity. They can also be
 CC administered to subjects undergoing gastrointestinal diagnostic
 CC investigation, particularly radiological or by magnetic resonance
 CC imaging.
 CC Extendins, components of Gila monster venom, have some sequence
 CC similarity to glucagon-like peptides (GLP). They are GLP agonists
 CC and have been suggested (US424286) for treatment of diabetes and
 CC prevention of hyperglycaemia.
 SQ Sequence 39 AA;

Query Match 91.8%; Score 212; DB 30; Length 39;
 Best Local Similarity 93.3%; Pred. No. 1.06e-10;
 Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 hsgdftsdlskqmeeeavrlfiewlknng 30
 QY 1 hgegtftsdlskqmeeeavrlfiewlknng 30

RESULT 11

ID R80545 standard; peptide; 39 AA.
 AC R80545;
 DT 27-FEB-1996 (first entry)
 DE Heloderma horridum extendin-3.
 KW Extendin-3; diabetes mellitus; hyperglycaemia; insulinotropic peptide.
 OS Heloderma horridum.
 PN US5424286-A.
 PD 13-JUN-1995.
 PR 24-MAY-1993; 066480.
 PR 24-MAY-1993; US-066480.
 PA (ENGJ/) ENG J.
 PI Eng J;
 DR WPI; 95-262627/34.
 PT Stimulating/inhibiting insulin release with extendin polypeptide(s) -
 PT for treating diabetes mellitus and preventing hyperglycaemia.
 PS Claim 5; Columns 13-14; 17pp; English.
 CC R80545 is Heloderma horridum extendin-3. It is an
 CC insulinotropic peptide, and can therefore be used in the treatment of
 CC diabetes mellitus (types I or II), and for the prevention of
 CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
 CC and insulin-(in)dependent mechanisms.
 SQ Sequence 39 AA;

Query Match 91.8%; Score 212; DB 14; Length 39;
 Best Local Similarity 93.3%; Pred. No. 1.06e-10;
 Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 hsgdftsdlskqmeeeavrlfiewlknng 30
 QY 1 hgegtftsdlskqmeeeavrlfiewlknng 30

RESULT 12
 ID W39309 standard; peptide; 30 AA.

AC W39309;
 DT 05-JUN-1998 (first entry)
 DE H. horridum extendin-4 peptide derivative #6.
 KW Extendin-3; extendin 4; diabetes; insulin; secretion; biosynthesis;
 KW glucagon reduction; hypoglycaemia; glucose; treatment.
 OS Heloderma horridum.
 FH Key Location/Qualifiers
 FT Modified_site 30 /note= "C-terminal amide"
 FT W09746584-A1.
 PN 11-DEC-1997.
 PD 05-JUN-1997; E02930.
 PE 13-SEP-1996; DE-037230.
 PR 03-JUN-1996; DE-022502.
 RA (BOEF) BOEHRINGER MANNHEIM GMBH.
 RA Goetze B, Goetze R, Hoffmann E;
 DR WPI: 98-042119/04.
 DT Truncated versions of extendin peptide(s) for treating diabetes -
 DT increase secretion and biosynthesis of insulin, but reduce those of
 DT glucagon, and do not induce hypoglycaemia
 PS Claim 2; Page 27; 150pp; English.
 PS Peptides W39303-W39420 are fragments of extendin-3 and extendin-4
 PS isolated from Heloderma horridum which are used in a novel method
 PS for the treatment of diabetes mellitus. These peptides can stimulate
 PS biosynthesis and secretion of insulin, but have the opposite effect on
 PS glucagon, and independent of this activity can increase peripheral
 PS glucagon utilisation. Extendin-3 and extendin-4 are only active when blood
 PS sugar levels are high, so they will not induce hypoglycaemia. Compared
 PS with glucagon-like peptide 1 (GLP1) and the known extendins, they are more
 PS active (effective at lower doses), more stable to degradation and
 PS metabolism and have a longer lasting effect. Truncated forms of this
 PS peptide can be made more economically than full length versions.
 SQ Sequence 30 AA;

Query Match 89.6%; Score 207; DB 29; Length 30;
 Best Local Similarity 96.6%; Pred. No. 3.03e-10;
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 hgegtfslskqmeeeavrlfiewlkg 29
 QY 1 hgegtfslskqmeeeavrlfiewlkg 29

RESULT 13
 ID W61772 standard; peptide; 28 AA.
 AC W61772;
 DT 29-MAR-1999 (first entry)
 DE Extendin-4 (1-28) amide for use in treating disorders related to food.
 KW Extendin; obesity; type II diabetes; eating disorders; cardiac disease;
 KW insulin resistance syndrome; elevated plasma glucose level; agonist.
 OS Heloderma suspectum
 FH Key Location/Qualifiers
 FT Modified_site 28 /note= "the C-terminal is in amide form"
 FT W09830231-A1.
 PN 16-JUL-1998.
 PD 07-JAN-1997; U00449.
 PE 14-NOV-1997; US-066029.
 PR 07-JAN-1997; US-034905.
 PR 08-AUG-1997; US-055404.
 PR 14-NOV-1997; US-065442.
 RA (AMYL-) AMYLIN PHARM INC.
 RA Beeley NRA, Bhavsar S, Prickett KS;
 DR WPI: 98-398796/34.
 DT Reducing food intake by administering extendins or their
 DT analogues - for treatment of e.g. obesity, type II diabetes,
 DT eating disorders and insulin resistance
 PS Claims 18, 26; Page 12; 214pp; English.
 PS The invention relates to a new method for treating disorders that
 PS are alleviated by reducing food intake, in particular obesity, type
 PS II diabetes, eating disorders, insulin resistance syndrome, elevated
 PS plasma glucose levels, or the risk of cardiac disease. The method
 PS comprises administering an extendin or an extendin agonist. The treatment

CC reduces appetite and lowers plasma lipid levels. It inhibits food
 CC consumption as effectively as amylin or cholecystokinin but has a much
 CC longer-lasting action (still effective after 6 hours in a mouse model).
 CC the present sequence is that of extendin-4 (1-28) amide which is one of
 CC the preferred compounds for use in the method.
 SQ Sequence 28 AA;

Query Match 88.7%; Score 205; DB 39; Length 28;
 Best Local Similarity 100.0%; Pred. No. 4.60e-10;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 hgegtfslskqmeeeavrlfiewlkn 28
 QY 1 hgegtfslskqmeeeavrlfiewlkn 28

RESULT 14
 ID W39368 standard; peptide; 30 AA.
 AC W39368;
 DT 05-JUN-1998 (first entry)
 DE H. horridum extendin-3 peptide derivative #11.
 DE Extendin-3; extendin 4; diabetes; insulin; secretion; biosynthesis;
 KW glucagon reduction; hypoglycaemia; glucose; treatment.
 OS Heloderma horridum.
 FH Key Location/Qualifiers
 FT Modified_site 30 /note= "C-terminal amide"
 FT W09746584-A1.
 PN 11-DEC-1997.
 PD 05-JUN-1997; E02930.
 PE 13-SEP-1996; DE-037230.
 PR 05-JUN-1996; DE-022502.
 RA (BOEF) BOEHRINGER MANNHEIM GMBH.
 RA Goetze B, Goetze R, Hoffmann E;
 DR WPI: 98-042119/04.
 DT Truncated versions of extendin peptide(s) for treating diabetes -
 DT increase secretion and biosynthesis of insulin, but reduce those of
 DT glucagon, and do not induce hypoglycaemia
 PS Claim 2; Page 27; 150pp; English.
 PS Peptides W39303-W39420 are fragments of extendin-3 and extendin-4
 PS isolated from Heloderma horridum which are used in a novel method
 PS for the treatment of diabetes mellitus. These peptides can stimulate
 PS biosynthesis and secretion of insulin, but have the opposite effect on
 PS glucagon, and independent of this activity can increase peripheral
 PS glucagon utilisation. Extendin-3 and extendin-4 are only active when blood
 PS sugar levels are high, so they will not induce hypoglycaemia. Compared
 PS with glucagon-like peptide 1 (GLP1) and the known extendins, they are more
 PS active (effective at lower doses), more stable to degradation and
 PS metabolism and have a longer lasting effect. Truncated forms of this
 PS peptide can be made more economically than full length versions.
 SQ Sequence 30 AA;

Query Match 88.7%; Score 205; DB 29; Length 30;
 Best Local Similarity 93.1%; Pred. No. 4.60e-10;
 Matches 27; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 hsdgtfslskqmeeeavrlfiewlkg 29
 QY 1 hgegtfslskqmeeeavrlfiewlkg 29

RESULT 15
 ID W39301 standard; peptide; 30 AA.
 AC W39301;
 DT 05-JUN-1998 (first entry)
 DE H. horridum extendin-3 peptide.
 KW Extendin-3; extendin 4; diabetes; insulin; secretion; biosynthesis;
 KW glucagon reduction; hypoglycaemia; glucose; treatment.
 OS Heloderma horridum.
 FH Key Location/Qualifiers
 FT Modified_site 30 /note= "This residue can be any amino acid except
 FT Gly"

PN WO9746584-A1.
 PD 11-DEC-1997.
 PF 05-JUN-1997; E02930.
 PR 13-SEP-1996; DE-037230.
 PR 05-JUN-1996; DE-022502.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Goeke B, Goeke R, Hoffmann E;
 DR WPI: 98-042119/04.
 PT Truncated versions of exendin peptide(s) for treating diabetes -
 PT increase secretion and biosynthesis of insulin, but reduce those of
 PT glucagon, and do not induce hypoglycaemia
 PS Claim 1; Page 3; 150pp; English.
 CC This peptide is a fragment of exendin-3 isolated from Heloderma
 CC horridum. This peptide and its salts, esters and derivatives can be
 CC used to treat diabetes mellitus. They stimulate biosynthesis and
 CC secretion of insulin, but have the opposite effect on glucagon, and
 CC independent of this activity can increase peripheral glucose utilisation.
 CC Exendin-3 and exendin-4 are only active when blood sugar levels are
 CC high, so they will not induce hypoglycaemia. Compared with glucagon-like
 CC peptide 1 (GLP1) and the known exendins, they are more active (effective
 CC at lower doses), more stable to degradation and metabolism and have a
 CC longer lasting effect. Truncated forms of this peptide can be made more
 CC economically than full length versions.
 SQ Sequence 30 AA;

Query Match 88.7%; Score 205; DB 29; Length 30;
 Best Local Similarity 93.1%; Pred.No. 4.60e-10;
 Matches 27; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 hsdgftsdlskgmeeeavrlfiewlkng 29
 |:::|||||
 QY 1 hgegtftsdlskgmeeeavrlfiewlkng 29

Search completed: Mon Oct 4 15:26:33 1999
 Job time : 19 secs.

W P S R L H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Oct 4 15:25:48 1999; MasPar time 6.68 Seconds
Global output not generated. 185.964 Million cell updates/sec

File: >MOHAM-312-CLAIM82B.PEP
Description: (1-31) from moham312177.pep
Perfect Score: 231
Sequence: 1 hgegtfslskmqeeavrlfiewlknggy 31

Scoring table: PAM 150
Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pirf60
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 34.253; Variance 64.514; scale 0.531

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	219	94.8	39	1 HWGH4G	extendin-4 - Gila mons	2.37e-25
2	212	91.8	39	1 HWGH3Z	extendin-3 - Mexican b	5.61e-24
3	127	55.0	31	2 S44472	glucagon G2 - North A	3.78e-08
4	124	53.7	101	1 GCFGB	glucagon precursor -	1.24e-07
5	122	52.8	30	2 S44473	glucagon-like peptide	2.73e-07
6	122	52.8	31	2 S44471	glucagon G1 - North A	2.73e-07
7	121	52.4	30	2 C61125	glucagon-like peptide	4.05e-07
8	121	52.4	30	2 B61125	glucagon-like peptide	4.05e-07
9	120	51.9	63	1 GCIDC	glucagon precursor -	5.99e-07
10	120	51.9	66	2 I51093	glucagon - chinook sa	5.99e-07
11	120	51.9	178	2 I51058	glucagon II precursor	5.99e-07
12	120	51.9	178	2 I51057	glucagon I precursor	5.99e-07
13	119	51.5	72	1 GCGXA	glucagon precursor -	8.84e-07
14	118	51.1	122	1 GCAF2	glucagon 2 precursor	1.30e-06
15	117	50.6	60	1 GCQNC	glucagon precursor -	1.92e-06
16	115	49.8	29	1 GCGF	glucagon - smaller sp	4.16e-06
17	115	49.8	158	1 GCPG	glucagon precursor -	4.16e-06
18	115	49.8	180	2 A57294	glucagon precursor -	4.16e-06
19	115	49.8	180	1 GCHY	glucagon precursor -	4.16e-06
20	115	49.8	180	1 GCRTDU	glucagon precursor -	4.16e-06
21	115	49.8	180	1 GCBO	glucagon precursor -	4.16e-06
22	115	49.8	180	1 GCPG	glucagon precursor -	4.16e-06
23	115	49.8	180	1 GCHU	glucagon precursor -	4.16e-06

24	115	49.8	180	1 GCRT	glucagon precursor -	4.16e-06
25	114	49.4	151	1 GCCH	glucagon precursor -	6.11e-06
26	114	49.4	206	2 I51301	proglucagon - chicken	6.11e-06
27	113	48.9	124	1 GCAF	glucagon 1 precursor	8.96e-06
28	110	47.6	29	1 GCBF	glucagon - Chinchilla	2.80e-05
29	108	46.8	29	2 S07211	glucagon - marbled el	2.96e-05
30	104	45.0	29	1 GCFLE	glucagon - European f	2.64e-04
31	104	45.0	29	2 A61135	glucagon - bigeye tun	2.64e-04
32	104	45.0	97	1 GCFIS	glucagon precursor -	2.64e-04
33	100	43.3	29	2 A91742	glucagon - Arabian ca	1.15e-03
34	100	43.3	29	2 A91741	glucagon - rabbit (te	1.15e-03
35	100	43.3	29	2 C39258	glucagon - common squ	1.15e-03
36	100	43.3	69	1 GCDG69	glucagon-69 - dog	1.15e-03
37	99	42.9	29	2 S39018	glucagon - bowfin	1.65e-03
38	99	42.9	29	1 GCEN	glucagon - elephantfi	1.65e-03
39	98	42.4	29	2 C60840	glucagon I - European	2.36e-03
40	96	41.6	29	1 GCOPIV	glucagon - North Amer	4.84e-03
41	96	41.6	29	2 A91740	glucagon - turkey (te	4.84e-03
42	95	41.1	29	1 A61583	glucagon - ostrich	6.91e-03
43	95	41.1	29	1 GCTTS	glucagon - slider tur	6.91e-03
44	95	41.1	29	1 GCDK	glucagon - duck	6.91e-03
45	91	39.4	36	1 GCFI	glucagon-36 - spotted	2.82e-02

ALIGNMENTS

RESULT 1 HWGH4G #type complete
ENTRY extendin-4 - Gila monster
TITLE #formal_name Heloderma suspectum #common_name Gila monster
ORGANISM 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
DATE 21-Nov-1997

ACCESSIONS A42486
REFERENCE A42486
#authors Eng, J.; Kleinman, W.A.; Singh, L.; Singh, G.; Raufman, J.P.
#journal J. Biol. Chem. (1992) 267:7402-7405
#title Isolation and characterization of extendin-4, an extendin-3 analogue, from Heloderma suspectum venom. Further evidence for an extendin receptor on dispersed acini from guinea pig pancreas

#cross-references MUID:92218391
#accession A42486
#molecule_type protein
#residues 1-39 #label ENG
COMMENT Extendin-4 does not stimulate amylase secretion by pancreatic acinar cells.

CLASSIFICATION #superfamily glucagon
KEYWORDS amidated carboxyl end; duplication; venom
FEATURE 39
#modified_site amidated carboxyl end (Ser) #status experimental
SUMMARY #length 39 #molecular-weight 4188 #checksum 9570
Query Match 94.8%; Score 219; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 2.37e-25;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 HGEFTFSLSKMQEEAVRLFIEWLKNGG 30
|||||
Qy 1 hgegtfslskmqeeavrlfiewlkngg 30
|||||

RESULT 2 HWGH3Z #type complete
ENTRY extendin-3 - Mexican beaded lizard
TITLE #formal_name Heloderma horridum #common_name Mexican beaded lizard
ORGANISM 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
DATE 21-Nov-1997
ACCESSIONS A23674
REFERENCE A23674
#authors Eng, J.; Andrews, P.C.; Kleinman, W.A.; Singh, L.; Raufman, J.P.

```

#Journal      J. Biol. Chem. (1990) 265:20259-20262
#Title        Purification and structure of Exendin-3, a new pancreatic
#cross-references MUID:91056067
#accession    A23674
#molecule_type protein
#residues     1-39 #label ENG
COMMENT       Exendins are venom components that are thought to bind to receptors
              for vasoactive intestinal peptide and/or secretin on pancreatic
              acinar cells and to activate adenylate cyclase, resulting in
              secretion of amylase.
CLASSIFICATION #superfamily glucagon
KEYWORDS       amidated carboxyl end; duplication; secretagogue; venom
FEATURE
39            #modified_site amidated carboxyl end (Ser) #status
              experimental
SUMMARY       #length 39 #molecular-weight 4204 #checksum 9591
Query Match   91.8%; Score 212; DB 1; Length 39;
Best Local Similarity 93.3%; Pred. No. 5.61e-24;
Matches       28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 1 HSDGFTSDLSKQMEEEAVRLFIEWLKNGG 30
|:::|||||:::|||||:::|||||:::|||||
QY 1 hgegtftsdlsgmeeeavrlfiewlknng 30
|:::|||||:::|||||:::|||||:::|||||

RESULT 3
ENTRY   S44472 #type complete
TITLE   glucagon G2 - North American paddleafish (Polyodon spathula)
ORGANISM #formal_name Polyodon spathula
DATE     19-Mar-1997 #sequence_revision 12-Dec-1997 #text_change
20-Mar-1998
ACCESSIONS S44472
REFERENCE   S44467
#authors   Nguyen, T.M.; Mommsen, T.P.; Mims, S.M.; Conlon, J.M.
#journal   Biochem. J. (1994) 300:339-345
#title     Characterization of insulins and proglucagon-derived peptides
              from a phylogenetically ancient fish, the paddlefish
              (Polyodon spathula).
#accession S44472
#molecule_type protein
#residues  1-31 #label NGU
#note      the sequence from Fig. 3 is inconsistent with that from
              Fig. 5 in having 29-Glu
SSIFICATION #superfamily glucagon
WORDS       carbohydrate metabolism; duplication; hormone; pancreas
FEATURE     1-31
#product glucagon G2 #status predicted #label GCN
SUMMARY     #length 31 #molecular-weight 3682 #checksum 7826
Query Match   55.0%; Score 127; DB 2; Length 31;
Best Local Similarity 55.2%; Pred. No. 3.78e-08;
Matches       16; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
Ph 1 HSCGMFTNDYSYLEEKSKEFVWLKNG 29
|:::|||||:::|||||:::|||||:::|||||
QY 1 hgegtftsdlsgmeeeavrlfiewlknng 29
|:::|||||:::|||||:::|||||:::|||||

RESULT 4
ENTRY   GCFGB #type fragments
TITLE   glucagon precursor - bullfrog (fragments)
ALTERNATE_NAMES oxynotomodulin
CONTAINS       glucagon; glucagon-36 (oxynotomodulin); glucagon-like peptide
              1; glucagon-like peptide 2
ORGANISM #formal_name Rana catesbeiana #common_name bullfrog
DATE     31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
20-Mar-1998
ACCESSIONS B28091; D28091
REFERENCE   A92730
#authors   Pollock, H.G.; Hamilton, J.W.; Rouse, J.B.; Ebner, K.E.;
              Rawitch, A.B.

```

```

#Journal      J. Biol. Chem. (1988) 263:9746-9751
#Title        Isolation of peptide hormones from the pancreas of the
              bullfrog (Rana catesbeiana). Amino acid sequences of
              pancreatic polypeptide, oxynotomodulin, and two
              glucagon-like peptides.
#cross-references MUID:88257102
#accession    B28091
#molecule_type protein
#residues     1-36 #label PO2
#accession    C28091
#molecule_type protein
#residues     37-68 #label PO1
#accession    D28091
#molecule_type protein
#residues     69-101 #label PO3
CLASSIFICATION #superfamily glucagon
KEYWORDS       carbohydrate metabolism; duplication; hormone; pancreas
FEATURE
1-36          #product glucagon-36 (oxynotomodulin) #status
              experimental #label G36\
1-29          #product glucagon #status predicted #label GCN\
37-67         #product glucagon-like peptide 1 #status experimental
              #label GL1\
69-101        #product glucagon-like peptide 2 #status experimental
              #label GL2
SUMMARY       #length 101 #checksum 9108
Query Match   53.7%; Score 124; DB 1; Length 101;
Best Local Similarity 51.7%; Pred. No. 1.24e-07;
Matches       15; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
Db 37 HADGFTSDMSYLEEKAKEFVDWLKNG 65
|:::|||||:::|||||:::|||||:::|||||
QY 1 hgegtftsdlsgmeeeavrlfiewlknng 29
|:::|||||:::|||||:::|||||:::|||||

RESULT 5
ENTRY   S44473 #type complete
TITLE   glucagon-like peptide - North American paddleafish (Polyodon
              spathula)
ORGANISM #formal_name Polyodon spathula
DATE     18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change
20-Mar-1998
ACCESSIONS S44473
REFERENCE   S44467
#authors   Nguyen, T.M.; Mommsen, T.P.; Mims, S.M.; Conlon, J.M.
#journal   Biochem. J. (1994) 300:339-345
#title     Characterization of insulins and proglucagon-derived peptides
              from a phylogenetically ancient fish, the paddlefish
              (Polyodon spathula).
#accession S44473
#molecule_type protein
#residues  1-30 #label NGU
CLASSIFICATION #superfamily glucagon
KEYWORDS       duplication; hormone; pancreas
FEATURE     1-30
#product glucagon-like peptide #status predicted #label
              MAT
SUMMARY       #length 30 #molecular-weight 3359 #checksum 5186
Query Match   52.8%; Score 122; DB 2; Length 30;
Best Local Similarity 55.2%; Pred. No. 2.73e-07;
Matches       16; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
Db 1 HADGTYTSDASFLQEQNARDFISWLKNG 29
|:::|||||:::|||||:::|||||:::|||||
QY 1 hgegtftsdlsgmeeeavrlfiewlknng 29
|:::|||||:::|||||:::|||||:::|||||

RESULT 6
ENTRY   S44471 #type complete
TITLE   glucagon G1 - North American paddleafish (Polyodon spathula)
ORGANISM #formal_name Polyodon spathula

```

```
#title      The primary structure of glucagon-like peptide but not
#cross-references MUID:91340068
#accession   B61125
#molecule_type protein
#residues    1-30 ##label CON
#CLASSIFICATION
#KEYWORDS    amidated carboxyl end; duplication
FEATURE
1-30        #product glucagon-like peptide #status experimental
              #label GLP\
              #modified_site amidated carboxyl end (Arg); #status
              predicted
SUMMARY
length 30 #molecular-weight 3376 #checksum 6092

Query Match          52.4%; Score 121; DB 2; Length 30;
Best Local Similarity 48.3%; Pred. No. 4.09e-07;
Matches 14; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Ddb 1 HAEGTTSDVSSYLQDAQAKDFITWLKSG 29
|::||::|||::|::||::|::|::|
QY 1 hgegtftsdlkqmeeeavrlfiwlnkg 29
|::||::|||::|::||::|::|::|

RESULT 9
ENTRY      #type fragments
TITLE      glucagon precursor - channel catfish (fragments)
ORGANISM   #formal_name Ictalurus punctatus #common_name channel catfish
DATE       31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
ACCESSIONS A05166; A05167
REFERENCE   A92514
AUTHORS    Andrews, P.C.; Ronner, P.
JOURNAL    J. Biol. Chem. (1985) 260:3910-3914
TITLE      Isolation and structures of glucagon and glucagon-like
            peptide from catfish pancreas.
#cross-references MUID:85157536
#accession   A05166
#molecule_type protein
#residues    1-29 ##label AND1
#accession   A05167
#molecule_type protein
#residues    30-63 ##label AND2
#CLASSIFICATION
#KEYWORDS    #superfamily glucagon
            carbohydrate metabolism; duplication; hormone; pancreas
FEATURE
1-29        #product glucagon #status experimental #label CGN\
30-63        #product glucagon-like peptide 1 #status experimental
            #label GLI
SUMMARY
length 63 #checksum 9366

Query Match          51.9%; Score 120; DB 1; Length 63;
Best Local Similarity 48.3%; Pred. No. 5.99e-07;
Matches 14; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

Ddb 30 HADGTYTSDVSSYLQDAQAKDFITWLKSG 58
|::||::|||::|::||::|::|::|
QY 1 hgegtftsdlkqmeeeavrlfiwlnkg 29
|::||::|||::|::||::|::|::|

RESULT 10
ENTRY      #type fragment
TITLE      glucagon - chinook salmon (fragment)
ORGANISM   #formal_name Oncorhynchus tshawytscha #common_name chinook
            salmon
DATE       13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
ACCESSIONS I51093
REFERENCE   I51093
AUTHORS    Irwin, D.M.; Wong, J.
JOURNAL    Mol. Endocrinol. (1995) 9:267-277
TITLE      Trout and chicken proglucagon: alternative splicing generates
```

#cross-references MUID:95295739
#accession I51093
#status preliminary; translated from GB/EMBL/DBDJ
#molecule_type mRNA
#residues 1-66 #label IRW
#cross-references EMBL:U19020; NID:g736366; PID:g736367

CLASSIFICATION #superfamily glucagon
KEYWORDS duplication
SUMMARY #length 66 #checksum 1440

Query Match 51.9%; Score 120; DB 2; Length 66;
Best Local Similarity 44.8%; Pred. No. 5.99e-07;
Matches 13; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

33 HADGTYSDVSTYLDQQAARDFVSLKSG 61
:::||||| ::::| :|::|
1 hgeftsdiskmeeeavrlfiewlknng 29

RESULT 11
ENTRY I51058 #type complete
TITLE glucagon I precursor - rainbow trout
ORGANISM #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Nov-1997
ACCESSIONS I51058; I51299; I51056; I51037; I51036; I51300
REFERENCE A55895
#authors Irwin, D.M.; Wong, J.
#journal Mol. Endocrinol. (1995) 9:267-277
#title Trout and chicken proglucagon: alternative splicing generates mRNA transcripts encoding glucagon-like peptide 2.
#cross-references MUID:95295739
#accession I51058
#status preliminary; translated from GB/EMBL/DBDJ
#molecule_type mRNA
#residues 1-178 #label IRW
#cross-references EMBL:U19917; NID:g736364; PID:g736365; GB:S78475;
#accession I51299
#status preliminary; translated from GB/EMBL/DBDJ
#molecule_type mRNA
#residues 52-53, 'X', 55-123 #label IR2
#cross-references GB:S78473; NID:g999382; PID:g999383
#accession I51056
#status preliminary; translated from GB/EMBL/DBDJ
#molecule_type mRNA
#residues 58-123 #label IR3
#cross-references EMBL:U19913; NID:g736360; PID:g736361
#accession I51037
#status preliminary; translated from GB/EMBL/DBDJ
#molecule_type DNA
#residues 'M', 114-144 #label IR4
#cross-references EMBL:U19919; NID:g736374; PID:g736377
#accession I51036
#status preliminary; translated from GB/EMBL/DBDJ
#molecule_type DNA
#residues 113-123 #label IR5
#cross-references EMBL:U19918; NID:g736373; PID:g736376

GENETICS 123/2
#introns #superfamily glucagon
CLASSIFICATION duplication
KEYWORDS #length 178 #molecular-weight 20034 #checksum 5250
SUMMARY

Query Match 51.9%; Score 120; DB 2; Length 178;
Best Local Similarity 44.8%; Pred. No. 5.99e-07;
Matches 13; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

Db 90 HADGTYSDVSTYLDQQAARDFVSLKSG 118
:::||||| ::::| :|::|
1 hgeftsdiskmeeeavrlfiewlknng 29

QY

RESULT 12
ENTRY I51057 #type complete
TITLE glucagon II precursor - rainbow trout
ORGANISM #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Nov-1997
ACCESSIONS I51057; I51039; I51038
REFERENCE A55895
#authors Irwin, D.M.; Wong, J.
#journal Mol. Endocrinol. (1995) 9:267-277
#title Trout and chicken proglucagon: alternative splicing generates mRNA transcripts encoding glucagon-like peptide 2.
#cross-references MUID:95295739
#accession I51057
#status preliminary; translated from GB/EMBL/DBDJ
#molecule_type mRNA
#residues 1-178 #label IRW
#cross-references EMBL:U19914; NID:g736362; PID:g736363
#accession I51039
#status preliminary; translated from GB/EMBL/DBDJ
#molecule_type DNA
#residues 113-144 #label IR2
#cross-references EMBL:U19916; NID:g736369; PID:g736372
#accession I51038
#status preliminary; translated from GB/EMBL/DBDJ
#molecule_type DNA
#residues 113-123 #label IR3
#cross-references EMBL:U19915; NID:g736368; PID:g736371

GENETICS 123/2
#introns #superfamily glucagon
CLASSIFICATION duplication
KEYWORDS #length 178 #molecular-weight 19998 #checksum 4464
SUMMARY

Query Match 51.9%; Score 120; DB 2; Length 178;
Best Local Similarity 44.8%; Pred. No. 5.99e-07;
Matches 13; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

Db 90 HADGTYSDVSTYLDQQAARDFVSLKSG 118
:::||||| ::::| :|::|
QY 1 hgeftsdiskmeeeavrlfiewlknng 29

RESULT 13
ENTRY GCGXA #type fragment
TITLE glucagon precursor - alligator gar (fragment)
ALTERNATE_NAMES oxyntomodulin
CONTAINS glucagon; glucagon-36 (oxyntomodulin); glucagon-like peptide 1
ORGANISM #formal_name Lepisosteus spatula #common_name alligator gar
DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 20-Mar-1998
ACCESSIONS S06339; S06871
REFERENCE S06339
#authors Pollock, H.G.; Kimmel, J.R.; Ebner, K.E.; Hamilton, J.W.; Rouse, J.B.; Lance, V.; Rawitch, A.B.
#journal Gen. Comp. Endocrinol. (1988) 69:133-140
#title Isolation of alligator gar (Lepisosteus spatula) glucagon, oxyntomodulin, and glucagon-like peptide: amino acid sequences of oxyntomodulin and glucagon-like peptide.

#cross-references MUID:88196798
#accession S06339
#molecule_type protein
#residues 1-36 #label POL
#accession S06871
#molecule_type protein
#residues 39-72 #label PO2
COMMENT X's at positions 37-38 represent a pair of basic amino acids forming a cleavage site.
CLASSIFICATION #superfamily glucagon
KEYWORDS carbohydrate metabolism; duplication; hormone; pancreas
FEATURE

1-36 #product glucagon-36 (oxyntomodulin) #status
experimental #label G36\
1-29 #product glucagon #status predicted #label GCN\
39-70 #product glucagon-like peptide 1 #status predicted
#label GLI
SUMMARY #length 72 #checksum 8055

Query Match 51.5%; Score 119; DB 1; Length 72;
Best Local Similarity 44.8%; Pred. No. 8.84e-07;
Matches 13; Conservative 10; Mismatches 6; Indels 0; Gaps 0;
Db 39 HADGTYTSDVSSYLQDQAARFVWLKQG 67
QY 1 hgegtftsdlsgmeeeavrlfiewlkg 29

RESULT 14 GCAR2 #type complete
ENTRY glucagon 2 precursor - American goosetish
TITLE glucagon; glucagon-like peptide 1
CONTAINS #formal_name Lophius americanus #common_name American
ORGANISM goosetish
3 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
20-Mar-1998

ACCESSIONS A05150
REFERENCE A05150
#authors Lund, P.K.; Goodman, R.H.; Montminy, M.R.; Dee, P.C.;
Habener, J.F.
#journal J. Biol. Chem. (1983) 258:3280-3284
#title Anglerfish islet pre-proglucagon II. Nucleotide and
corresponding amino acid sequence of the cDNA.
#cross-references MUID:83135785

#accession A05150
#molecule_type mRNA
#residues 1-122 #label LUN
#cross-references GB:T09933; NID:g213352; PID:g213353
CLASSIFICATION #superfamily glucagon
KEYWORDS carbohydrate metabolism; duplication; hormone; pancreas
FEATURE

1-21 #domain signal sequence #status predicted #label SIG\
22-122 #product proglucagon 2 #status predicted #label PGC2\
52-80 #product glucagon #status predicted #label GCN\
89-119 #product glucagon-like peptide 1 #status predicted
#label GLI
SUMMARY #length 122 #molecular-weight 14171 #checksum 7194

Query Match 51.1%; Score 118; DB 1; Length 122;
Best Local Similarity 44.8%; Pred. No. 1.30e-06;
Matches 13; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

89 HADGTYTSDVSSYLQDQAARFVWLKAG 117
QY 1 hgegtftsdlsgmeeeavrlfiewlkg 29

RESULT 15 GCOMC #type fragments
ENTRY glucagon precursor - coho salmon (fragments)
TITLE glucagon; glucagon-like peptide 1
CONTAINS #formal_name Oncorhynchus kisutch #common_name coho salmon
ORGANISM 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change
20-Mar-1998

ACCESSIONS JP0103; JP0104
REFERENCE A94232
#authors Plisetzkaya, E.; Pollock, H.G.; Rouse, J.B.; Hamilton, J.W.;
Kimmel, J.R.; Gorbman, A.
#journal Regul. Pept. (1986) 14:57-67
#title Isolation and structure of coho salmon (Oncorhynchus kisutch)
glucagon and glucagon-like peptide.
#cross-references MUID:86234328
#accession JP0103
#molecule_type protein
#residues 1-29;30-60 #label PLI

CLASSIFICATION #superfamily glucagon
KEYWORDS carbohydrate metabolism; duplication; hormone; pancreas
FEATURE
1-29 #product glucagon #status experimental #label GCN\
30-60 #product glucagon-like peptide 1 #status experimental
#label GLI
SUMMARY #length 60 #checksum 8993

Query Match 50.6%; Score 117; DB 1; Length 60;
Best Local Similarity 41.4%; Pred. No. 1.92e-06;
Matches 12; Conservative 12; Mismatches 5; Indels 0; Gaps 0;
Db 30 HADGTYTSDVSSYLQDQAARFVWLKSG 58
QY 1 hgegtftsdlsgmeeeavrlfiewlkg 29

Search completed: Mon Oct 4 15:25:57 1999
Job time : 9 secs.

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Oct 4 15:24:52 1999; MasPar time 3.93 Seconds

abular output not generated. 222.856 Million cell updates/sec

File: >MOHAM-312-CLAIM82B.PEP
Description: (1-31) from moham312177.pep
Perfect Score: 231
Sequence: 1 hgegttsdlskqmeeeavrfiewlknngy 31

Scoring table: PAM 150
Gap 11

Searched: 77977 seqs, 28269293 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot37
1:swissprot

Statistics: Mean 35.168; Variance 58.938; scale 0.597

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	219	94.8	87	1	EXE4_HELVSU	1.37e-28
2	212	91.8	39	1	EXE3_HELHO	4.69e-27
3	124	53.7	103	1	GLUC_RANCA	7.61e-09
4	121	52.4	30	1	GLUM_RANCA	2.82e-08
5	121	52.4	71	1	GLUC_ICTPU	2.82e-08
6	119	51.5	78	1	GLUC_LEPSP	6.70e-08
7	118	51.1	122	1	GLUC_LOPKM	1.03e-07
8	117	50.6	68	1	GLUC_ONCKI	1.59e-07
9	115	49.8	29	1	GLUC_SCYCA	3.73e-07
10	115	49.8	158	1	GLUC_PIG	3.73e-07
11	115	49.8	180	1	GLUC_OCTDE	3.73e-07
12	115	49.8	180	1	GLUC_MOUSE	3.73e-07
13	115	49.8	180	1	GLUC_BOVIN	3.73e-07
14	115	49.8	180	1	GLUC_HUMAN	3.73e-07
15	115	49.8	180	1	GLUC_CAVPO	3.73e-07
16	115	49.8	180	1	GLUC_MESAU	3.73e-07
17	115	49.8	180	1	GLUC_CARAU	3.73e-07
18	114	49.4	121	1	GLUC_CHICK	5.71e-07
19	114	49.4	151	1	GLUC_LOPKM	8.72e-07
20	113	48.9	124	1	GLUC_CHIBR	3.08e-06
21	110	47.6	29	1	GLUC_TORMA	7.10e-06
22	108	46.8	29	1	GLUC2_ORENI	2.45e-05
23	105	45.5	33	1	GLUC2_ORENI	2.45e-05

24	104	45.0	29	1	GLUC_PLAFE	GLUCAGON.	3.70e-05
25	104	45.0	96	1	GLUC_MYOSC	GLUCAGON.	3.70e-05
26	100	43.3	29	1	GLUC_RABIT	GLUCAGON.	1.87e-04
27	100	43.3	69	1	GLUC_CANFA	GLUCAGON.	1.87e-04
28	99	42.9	29	1	GLUC_CALMI	GLUCAGON.	2.79e-04
29	99	42.9	75	1	GLUC_AMICA	GLUCAGON.	2.79e-04
30	98	42.4	36	1	GLUL_ORENI	GLUCAGON.	4.16e-04
31	96	41.6	29	1	GLUC_DIDMA	GLUCAGON.	9.17e-04
32	95	41.1	29	1	GLUC_ANAPL	GLUCAGON.	1.36e-03
33	91	39.4	36	1	GLUC_HYDCO	GLUCAGON.	6.40e-03
34	88	38.1	2127	1	RRPL_RABVS	RNA POLYMERASE BETA SU	2.00e-02
35	88	38.1	2142	1	RRPL_RABVP	RNA POLYMERASE BETA SU	2.00e-02
36	83	35.9	406	1	YBDN_ECOLI	HYPOTHETICAL 47.8 KD P	1.27e-01
37	81	35.1	658	1	UVRB_HELPY	EXONUCLEASE ABC SUBUN	2.61e-01
38	80	34.6	170	1	VIP_HUMAN	VASOACTIVE INTESTINAL	3.73e-01
39	79	34.2	42	1	GIP_PIG	GASTRIC INHIBITORY POL	5.31e-01
40	79	34.2	42	1	GIP_BOVIN	GASTRIC INHIBITORY POL	5.31e-01
41	79	34.2	144	1	GIP_MOUSE	GASTRIC INHIBITORY POL	5.31e-01
42	79	34.2	144	1	GIP_MOUSE	GASTRIC INHIBITORY POL	5.31e-01
43	79	34.2	153	1	GIP_HUMAN	GASTRIC INHIBITORY POL	5.31e-01
44	78	33.8	134	1	SECR_RAT	SECRETIN PRECURSOR.	7.53e-01
45	78	33.8	456	1	ENO_MYCPN	ENOLASE (EC 4.2.1.11)	7.53e-01

ALIGNMENTS

RESULT 1
ID EXE4_HELVSU STANDARD; PRT; 87 AA.
AC F26349;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE EXENDIN-4 PRECURSOR.
OS HELODERMA SUSPECTUM (GILA MONSTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;
OC SCLEROGLOSSA; ANGIOMORPHA; HELODERMATIDAE; HELODERMA.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN Y.E., DRUCKER D.J.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 48-86.
RC TISSUE=VENOM;
RX MEDLINE; 92218391.
RA ENG J., KLEINMAN W.A., SINGH L., SINGH G., RAUFMAN J.-P.;
RT "Isolation and characterization of exendin-4, an exendin-3 analogue,
RT from Heloderma suspectum venom. Further evidence for an exendin
RT receptor on dispersed acini from guinea pig pancreas.";
RL J. BIOL. CHEM. 267:7402-7405(1992).
CC -!- FUNCTION: HAS A VIP/SECRETIN-LIKE BIOLOGICAL ACTIVITY. INTERACTS
CC WITH THE EXENDIN RECEPTOR.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC EMBL; U77613; G1916067; -.
CC PIR; A42486; HWGH4G.
DR PROSITE; PS00260; GLUCAGON; 1.
DR PFAM; PF00123; hormone2; 1.
KW GLUCAGON FAMILY; VENOM; AMIDATION; SIGNAL.
FT SIGNAL 1 23 POTENTIAL.
FT PEPTIDE 48 86 EXENDIN-4.
FT MOD_RES 86 86 AMIDATION (G-87 PROVIDE AMIDE GROUP).
SQ SEQUENCE 87 AA; 9479 MW; 6C1A8FD5 CRC32;
Query Match 94.8%; Score 219; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.37e-28;

```

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 48 HGEFTSDLSKQMEEEAVRLFIEWLKNG 77
QY 1 hgeftsdlskqmeeeavrlfiewlknng 30

RESULT 2
ID EXE3_HELHO STANDARD; PRT; 39 AA.
AC P20394;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE EXENDIN-3.
OS HELODERMA HORRIDUM HORRIDUM (MEXICAN BEADED LIZARD).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;
SC SCLEROGLOSSA; ANGUIMORPHA; HELODERMATIDAE; HELODERMA.
[1]
AP SEQUENCE.
RC TISSUE=VENOM;
RX MEDLINE; 91056067.
RA ENG J., ANDREW P.C., KLEINMAN W.A., SINGH L., RAUFMAN J.-P.;
RT "Purification and structure of exendin-3, a new pancreatic
secretagogue isolated from Heloderma horridum venom.";
RL J. BIOL. CHEM. 265:20259-20262(1990).
CC -!- FUNCTION: HAS A VIP/SECRETIN-LIKE BIOLOGICAL ACTIVITY. INTERACTS
CC WITH THE EXENDIN RECEPTOR.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR PIR; A23674; HWH32.
DR PROSITE; PS00260; GLUCAGON; 1.
DR PFAM; PF00123; hormone2; 1.
DR HSP; P01274; IGCN
KW GLUCAGON FAMILY; VENOM; AMIDATION.
FT MOD_RES 39
SQ SEQUENCE 39 AA; 4204 MW; AB598FD3 CRC32;

Query Match 91.8%; Score 212; DB 1; Length 39;
Best Local Similarity 93.3%; Pred. No. 4.69e-27;
Matches 28; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 HSDGTTSDLSKQMEEEAVRLFIEWLKNG 30
QY 1 hgeftsdlskqmeeeavrlfiewlknng 30

ULT 3
GLUC_RANCA STANDARD; PRT; 103 AA.
AC P15438; P15439; P15440;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE GLUCAGON PRECURSOR (FRAGMENTS).
OS RANA CATESBEIANA (BULL FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
SC NEOBATRACHIA; RANOIDEA; RANIDAE; RANINAE; RANA.
[1]
AP SEQUENCE.
RC TISSUE=PANCREAS;
RX MEDLINE; 88257102.
RA POLLOCK H.G., HAMILTON J.W., ROUSE J.B., EBNER K.E., RAWITCH A.B.;
RT "Isolation of peptide hormones from the pancreas of the bullfrog
(Rana catesbeiana). Amino acid sequences of pancreatic polypeptide,
oxyntomodulin, and two glucagon-like peptides.";
RL J. BIOL. CHEM. 263:9746-9751(1988).
CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC THE BLOOD SUGAR LEVEL.
CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC -!- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH OTHER SPECIES
CC SEQUENCES.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR PIR; B28091; GCGFB.
DR PROSITE; PS00260; GLUCAGON; 3.

```

```

DR PFAM; PF00123; hormone2; 3.
DR HSP; P01274; IGCN.
KW GLUCAGON FAMILY; HORMONE.
FT PEPTIDE 1 29
FT PEPTIDE 1 36
FT PEPTIDE 39 70
FT NON_CONS 70 71
FT PEPTIDE 71 103
SQ SEQUENCE 103 AA; 11719 MW; D43EDFC9 CRC32;

Query Match 53.7%; Score 124; DB 1; Length 103;
Best Local Similarity 51.7%; Pred. No. 7.61e-09;
Matches 15; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Db 39 HADGTTSDMSYLEEAKAEFVDWLKNG 67
QY 1 hgeftsdlskqmeeeavrlfiewlknng 29

RESULT 4
ID GLUC_ANGAN STANDARD; PRT; 30 AA.
AC P41521;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE GLUCAGON-LIKE PEPTIDE (GLP).
OS ANGUILLA ANGUILLA (EUROPEAN FRESHWATER EEL), AND
OS ANGUILLA ROSTRATA (AMERICAN EEL).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
SC TELEOSTEI; ANGUILLIFORMES; ANGUILLIDAE; ANGUILLA.
[1]
AP SEQUENCE.
RC TISSUE=PANCREAS;
RX MEDLINE; 91340068.
RA CONLON J.M., ANDREWS P.C., THIM L., MOON T.W.;
RT "The primary structure of glucagon-like peptide but not insulin has
been conserved between the American eel, Anguilla rostrata and the
European eel, Anguilla anguilla.";
RL GEN. COMP. ENDOCRINOL. 82:23-32(1991).
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR PIR; B61125; B61125.
DR PIR; C61125; C61125.
DR PROSITE; PS00260; GLUCAGON; 1.
DR PFAM; PF00123; hormone2; 1.
DR HSP; P01274; IGCN.
KW GLUCAGON FAMILY; AMIDATION.
FT MOD_RES 30
SQ SEQUENCE 30 AA; 3376 MW; 27E8C37D CRC32;

Query Match 52.4%; Score 121; DB 1; Length 30;
Best Local Similarity 48.3%; Pred. No. 2.82e-08;
Matches 14; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Db 1 HAGTTSVSSYLQDAKEFVSWLKTG 29
QY 1 hgeftsdlskqmeeeavrlfiewlknng 29

RESULT 5
ID GLUC_ICTPU STANDARD; PRT; 71 AA.
AC P04093;
DT 01-NOV-1986 (REL. 03, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
DE GLUCAGON PRECURSOR (FRAGMENT).
OS ICTALURUS PUNCTATUS (CHANNEL CATFISH).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
SC TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; SILURIFORMES; ICTALURIDAE;
[1]
AP SEQUENCE.
RC TISSUE=PANCREAS;
RX MEDLINE; 87156787.

```

RA HOSEIN N.M., MAHREHOLZ A.M., ANDREWS P.C., GURD R.S.;
 RT "Biological activities of catfish glucagon and glucagon-like
 RL peptide";
 RN BIOCHEM. BIOPHYS. RES. COMMUN. 143:97-92(1987).
 RN [2]
 RP SEQUENCE.
 RC TISSUE-PANCREAS;
 RX MEDLINE; 85157536.
 RA ANDREWS P.C., RONNER P.;
 RT "Isolation and structures of glucagon and glucagon-like peptide from
 RL catfish pancreas";
 RN J. BIOL. CHEM. 260:3910-3914(1985).
 CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -1- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH AMERICAN
 CC GOOSEFISH SEQUENCES.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 CC PIR; A05166; GCIDC.
 DR PROSITE; PS00260; GLUCAGON; 2.
 DR PFAM; PF00123; hormone2; 2.
 DR HSP; P01274; IGCN.
 DR GLUCAGON FAMILY; HORMONE.
 DR NON_TER 1 1
 FT PEPTIDE 1 29 GLUCAGON.
 FT PEPTIDE 38 71 GLUCAGON-LIKE PEPTIDE.
 FT CONFLICT 53 53 E -> D (IN REF. 2).
 FT NON_TER 71 71
 SQ SEQUENCE 71 AA; 8173 MW; C49ED93A CRC32;
 Query Match 52.4%; Score 121; DB 1; Length 71;
 Best Local Similarity 51.7%; Pred. No. 2.82e-08;
 Matches 15; Conservative 9; Mismatches 5; Indels 0; Gaps 0;
 Db 38 HADGTYTSDYSSYLQQAQAKFFITWLKSG 66
 QV 1 hgegtfslskmqeeavrlfiewlknk 29
 RESULT 6
 ID GLUC_LEPSP STANDARD; PRT; 78 AA.
 AC P09566;
 DT 01-MAR-1989 (REL. 10, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE GLUCAGON PRECURSOR (FRAGMENT).
 OS LEPISTOSTEUS SPATULA (ALLIGATOR GAR) (ATFACOSTEUS SPATULA).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 CC SEMIONOTIFORMES; LEPISTOSTEIDAE; LEPISTOSTEUS.
 CC [1]
 CC SEQUENCE OF 1-36 AND 45-78.
 CC TISSUE-PANCREAS;
 CC MEDLINE; 88196798.
 RA POLLOCK H.G., KIMMEL J.R., EBERNER K.E., HAMILTON J.W., ROUSE J.B.,
 RA LANCE V., RAWITCH A.B.;
 RT "Isolation and structures of alligator gar (Lepisosteus spatula) glucagon,
 RL oxyntomodulin, and glucagon-like peptide: amino acid sequences of
 RT oxyntomodulin and glucagon-like peptide";
 RL GEN. COMP. ENDOCRINOL. 69:133-140(1988).
 RN [2]
 RP PRELIMINARY SEQUENCE OF 1-29.
 RC TISSUE-PANCREAS;
 RX MEDLINE; 88030594.
 RA POLLOCK H.G., KIMMEL J.R., ROUSE J.B., EBERNER K.E.,
 RA LANCE V., RAWITCH A.B.;
 RT "Isolation and structures of alligator gar (Lepisosteus spatula)
 RL insulin and pancreatic polypeptide";
 RL GEN. COMP. ENDOCRINOL. 67:375-382(1987).
 CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.

CC -1- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH AMERICAN
 CC GOOSEFISH SEQUENCES.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PIR; S06339; GCGXA.
 DR PROSITE; PS00260; GLUCAGON; 2.
 DR PFAM; PF00123; hormone2; 2.
 DR HSP; P01274; IGCN.
 DR GLUCAGON FAMILY; HORMONE.
 DR PEPTIDE 1 29 GLUCAGON.
 FT PEPTIDE 1 36 GLUCAGON-36 (OXYNTOMODULIN).
 FT PEPTIDE 45 78 GLUCAGON-LIKE PEPTIDE.
 SQ SEQUENCE 78 AA; 8990 MW; 509ED9D3 CRC32;
 Query Match 51.5%; Score 119; DB 1; Length 78;
 Best Local Similarity 44.8%; Pred. No. 6.70e-08;
 Matches 13; Conservative 10; Mismatches 6; Indels 0; Gaps 0;
 Db 45 HADGTYTSDYSSYLQQAQAKFFITWLKSG 73.
 QV 1 hgegtfslskmqeeavrlfiewlknk 29
 RESULT 7
 ID GLU2_LOPAM STANDARD; PRT; 122 AA.
 AC P04092;
 DT 01-NOV-1986 (REL. 03, CREATED)
 DT 01-NOV-1986 (REL. 03, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE GLUCAGON II PRECURSOR.
 OS LOPHIUS AMERICANUS (AMERICAN GOOSEFISH) (ANGLERFISH).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 CC TELEOSTEI; EUTELEOSTEI; PARACANTHOPTERYGII; LOPHIIFORMES; LOPHIIDAE;
 CC LOPHIUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 83135785.
 RA LUND P.K., GOODMAN R.H., MONTMINY M.R., DEE P.C., HABENER J.F.;
 RT "Anglerfish islet pre-proglucagon II. Nucleotide and corresponding
 RL amino acid sequence of the cDNA";
 RL J. BIOL. CHEM. 258:3280-3284(1983).
 RN [2]
 RP PROCESSING.
 RX MEDLINE; 86286913.
 RA NOE B.D., ANDREWS P.C.;
 RT "Specific glucagon-related peptides isolated from anglerfish islets
 RL are metabolic cleavage products of (pre)proglucagon-II";
 RL PEPTIDES 7:331-339(1986).
 CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V00632; G64022;
 DR EMBL; J00933; G213353;
 DR PIR; A05150; GCAF2.
 DR PROSITE; PS00260; GLUCAGON; 2.
 DR PFAM; PF00123; hormone2; 2.
 DR HSP; P01274; IGCN.
 DR GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
 DR SIGNAL 1 21
 FT PEPTIDE 22 49 GRP (GLICENTINE RELATED POLYPEPTIDE).
 FT PEPTIDE 52 80 GLUCAGON II.
 FT PEPTIDE 89 119 GLUCAGON-LIKE PEPTIDE II.
 SQ SEQUENCE 122 AA; 14171 MW; DFE63061 CRC32;

THE
CONSTITUTION
OF THE
UNITED STATES
OF AMERICA

"X-ray analysis of glucagon and its relationship to receptor

"X-ray analysis of glucagon and its relationship to receptor

NATURE 257:751-757(1975).

1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES THE BLOOD SUGAR LEVEL.

1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.

1- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH HUMAN SEQUENCE.

1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.

PIR: A01540; GCPG.

PDB: 1GCN; 30-SEP-83.

PROSITE: PS00260; GLUCAGON; 3.

PFAM: PF00123; hormone2; 3.

GLUCAGON FAMILY: HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; 3D-STRUCTURE.

FT NON_TER 1 1 GRPP (GLICENTINE RELATED POLYPEPTIDE).

FT PEPTIDE 1 30 GLUCAGON.

FT PEPTIDE 33 61

FT PEPTIDE 78 107 GLUCAGON-LIKE PEPTIDE 1.

FT PEPTIDE 126 158 GLUCAGON-LIKE PEPTIDE 2.

FT HELIX 39 42

FT TURN 43 45

FT HELIX 46 55

TURN 56 57

SEQUENCE 136 AA; 18212 MW; 9FBC1BFE CRC32;

Query Match 49.8%; Score 115; DB 1; Length 158;
Best Local Similarity 55.2%; Pred. No. 3.73e-07;
Matches 16; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Db 78 HAEGFTSDVSYLGGQAAKEFIWLKVG 106
1 hgegtftsdlkqmeeeavrlfiewlkg 29
1 hgegtftsdlkqmeeeavrlfiewlkg 29

RESULT 11

ID	GLUC_RAT	STANDARD;	PRT;	180 AA.
AC	P06883;			
DT	01-JAN-1988 (REL. 06, CREATED)			
DT	01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)			
DT	01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)			
DE	GLUCAGON PRECURSOR.			
GN	GCG.			
OS	RATTUS NORVEGICUS (RAT).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.			
[1]				
RP	SEQUENCE FROM N.A.			
EX	MEDLINE; 85054853.			
RA	HEINRICH G., GROS P., HABENER J.F.;			
RT	"Glucagon gene sequence. Four of six exons encode separate functional domains of rat pre-proglucagon."			
J. BIOL. CHEM.	259:14082-14087(1984).			
[2]				
RP	SEQUENCE FROM N.A.			
EX	MEDLINE; 85051023.			
RA	HEINRICH G., GROS P., LUND P.K., BENTLEY R.C., HABENER J.F.;			
RT	"Pre-proglucagon messenger ribonucleic acid; nucleotide and encoded amino acid sequences of the rat pancreatic complementary deoxyribonucleic acid."			
RL	ENDOCRINOLOGY 115:2176-2181(1984).			
[3]				
RP	SEQUENCE FROM N.A.			
EX	MEDLINE; 86304324.			
RA	MOJUSOV S., HEINRICH G., WILSON I.B., RAVAZZOLA M., ORCI L.,			
RA	HABENER J.F.;			
RT	"Preproglucagon gene expression in pancreas and intestine diversifies at the level of post-translational processing."			
RL	J. BIOL. CHEM. 261:11880-11889(1986).			
1- FUNCTION:	PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES THE BLOOD SUGAR LEVEL.			
1- INDUCTION:	PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.			
1- SIMILARITY:	BELONGS TO THE GLUCAGON FAMILY.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; K02813; G204370; .

DR EMBL; K02809; G204370; JOINED.

DR EMBL; K02810; G204370; JOINED.

DR EMBL; K02811; G204370; JOINED.

DR EMBL; K02812; G204370; JOINED.

DR PIR; A22655; GCET.

DR PIR; A44198; A44198.

DR PROSITE; PS00260; GLUCAGON; 4.

DR PFAM; PF00123; hormone2; 3.

DR HSSP; P01274; 1GCN.

KW GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.

FT SIGNAL 1 20

FT PEPTIDE 21 50 GRPP (GLICENTINE RELATED POLYPEPTIDE).

FT PEPTIDE 53 81 GLUCAGON.

FT PEPTIDE 92 128 GLUCAGON-LIKE PEPTIDE 1.

FT PEPTIDE 146 178 GLUCAGON-LIKE PEPTIDE 2.

SQ SEQUENCE 180 AA; 20846 MW; 355C3843 CRC32;

Query Match 49.8%; Score 115; DB 1; Length 180;
Best Local Similarity 55.2%; Pred. No. 3.73e-07;
Matches 16; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Db 98 HAEGFTSDVSYLGGQAAKEFIWLKVG 126
1 hgegtftsdlkqmeeeavrlfiewlkg 29
1 hgegtftsdlkqmeeeavrlfiewlkg 29

RESULT 12

ID	GLUC_OCTDE	STANDARD;	PRT;	180 AA.
AC	P22890;			
DT	01-AUG-1991 (REL. 19, CREATED)			
DT	01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)			
DT	01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)			
DE	GLUCAGON PRECURSOR.			
OS	OCTODON DEGUS (DEGU).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	RODENTIA; HYSTRICOGNATHI; OCTODONTIDAE; OCTODON.			
[1]				
RP	SEQUENCE FROM N.A.			
EX	MEDLINE; 91155952.			
RA	NISHI M., STEINER D.F.;			
RT	"Cloning of complementary DNAs encoding islet amyloid polypeptide, insulin, and glucagon precursors from a New World rodent, the degu, Octodon degus."			
RL	MOL. ENDOCRINOL. 4:1192-1198(1990).			
1- FUNCTION:	PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES THE BLOOD SUGAR LEVEL.			
1- INDUCTION:	PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.			
1- SIMILARITY:	BELONGS TO THE GLUCAGON FAMILY.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M57688; G202468; .

DR PIR; C36118; GCETDU.

DR PROSITE; PS00260; GLUCAGON; 4.

DR PFAM; PF00123; hormone2; 3.

DR HSSP; P01274; 1GCN.


```

OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88330860.
RT DRUCKER D.J., ASA S.;
RA "Glucagon gene expression in vertebrate brain.";
RL J. BIOL. CHEM. 263:13475-13478(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86259053.
RT WHITE J.W., SAUNDERS G.F.;
RA "Structure of the human glucagon gene.";
RL NUCLEIC ACIDS RES. 14:4719-4730(1986).
RN [3]

```

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Oct 4 15:25:15 1999; MasPar time 10.69 Seconds
158.226 Million cell updates/sec

ular output not generated.

File: >MOHAM-312-CLAIM82B.PEP
Description: (1-31) from moham312177.pep
Perfect Score: 231
Sequence: 1 hgegtfslskmqeeavrlfiewlknngy 31

Scoring table: PAM 150
Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrmb19
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp_mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp_unclassified
13:sp-vertebrate 14:sp_virus

Statistics: Mean 33.531; Variance 62.579; scale 0.536

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match	Length	ID	Description	Pred. No.
1	146	63.2	266	13	PROGLUCAGON I.	9.94e-12
2	143	61.9	219	13	PROGLUCAGON II.	3.51e-11
3	120	51.9	66	13	GLUCAGON (FRAGMENT).	4.25e-07
4	120	51.9	72	13	PROGLUCAGON (FRAGMENT)	4.25e-07
5	120	51.9	72	13	PROGLUCAGON (FRAGMENT)	4.25e-07
6	120	51.9	178	13	PROGLUCAGON I.	4.25e-07
7	120	51.9	178	13	PROGLUCAGON II.	4.25e-07
8	114	49.4	206	13	PROGLUCAGON.	4.51e-06
9	108	46.8	149	13	PROGLUCAGON.	4.58e-05
10	108	46.8	204	13	PROGLUCAGON.	4.58e-05
11	95	41.1	2127	14	L PROTEIN, RNA DEPENDENT	5.81e-03
12	86	37.2	552	9	REPLICASE.	1.40e-01
13	81	35.1	844	5	HYPOTHETICAL 95.1 KD P	7.64e-01
14	78	33.8	1134	10	HYPOTHETICAL 127.3 KD	2.06e-00
15	77	33.3	414	5	R05H5.4 PROTEIN.	2.85e-00
16	77	33.3	562	10	FIN21.3.	2.85e-00
17	76	32.9	300	1	300AA LONG HYPOTHETICAL	3.93e-00
18	76	32.9	1086	4	NAD(P) TRANSHYDROGENAS	3.93e-00
19	76	32.9	1086	11	NICOTINAMIDE NUCLEOTID	3.93e-00
20	76	32.9	1086	4	NICOTINAMIDE NUCLEOTID	3.93e-00

21 75 32.5 324 4 Q16107
22 75 32.5 333 2 Q48538
23 75 32.5 502 5 Q22770
24 75 32.5 664 2 Q59339
25 75 32.5 666 11 Q88807
26 75 32.5 666 11 Q35117
27 75 32.5 1015 2 Q67146
28 75 32.5 1911 5 Q02038
29 74 32.0 283 2 Q46980
30 74 32.0 557 10 Q65538
31 74 32.0 605 3 Q00292
32 73 31.6 337 5 Q61798
33 73 31.6 418 3 Q14418
34 73 31.6 440 5 Q27456
35 73 31.6 454 5 Q27457
36 73 31.6 497 5 Q25615
37 73 31.6 970 10 Q80790
38 73 31.6 1041 5 Q18031
39 73 31.6 1141 3 Q74998
40 73 31.6 1352 1 Q58837
41 72 31.2 253 2 Q0547
42 72 31.2 839 5 Q62002
43 72 31.2 1106 4 Q15034
44 71 30.7 453 2 P72834
45 71 30.7 599 8 Q47430

C6.1A.
INTEGRASE/RECOMBINASE. 5.41e+00
T25B9.7 PROTEIN. 5.41e+00
ADENYLYL-SULPHATE REDU 5.41e+00
PEPTIDYLARGININE DEIMI 5.41e+00
PEPTIDYLARGININE DEIMI 5.41e+00
FORMATE DEHYDROGENASE 5.41e+00
VOLTAGE-GATED CALCIUM 5.41e+00
HOMOPROTocatechuate DY 7.42e+00
HYPOTHETICAL 61.6 KD P 7.42e+00
DIHYDROBODIN OXIDASE. 7.42e+00
C33E10.8 PROTEIN. 1.02e+01
CARBOXYPEPTIDASE. 1.02e+01
MICROFILARIAL CHITINAS 1.02e+01
MICROFILARIAL CHITINAS 1.02e+01
CHITINASE. 1.02e+01
REVERSE-TRANSCRIPTASE- 1.02e+01
NAD(P) TRANSHYDROGENAS 1.02e+01
NITRITE REDUCTASE (EC 1.02e+01
1352AA LONG HYPOTHETIC 1.02e+01
ESTERASE. 1.39e+01
INTERMEDIATE FILAMENT 1.39e+01
K1AA0318 (FRAGMENT). 1.39e+01
UDP-GLUCOSE DEHYDROGEN 1.89e+01
NADH DEHYDROGENASE SUB 1.89e+01

ALIGNMENTS

RESULT 1
ID Q42143 PRELIMINARY; PRT; 266 AA.
AC Q42143;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PROGLUCAGON I.
OC XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97368292.
RA IRWIN D.M., SATKUNARAJAH M., WEN Y., BRUBAKER P.L., PEDERSON R.A.,
RA WHEELER M.B.;
RT "The xenopus proglucagon gene encodes novel GLP-1-like peptides with
insulinotropic properties."
RL PROC. NATL. ACAD. SCI. U.S.A. 94:7915-7920(1997).
DR EMBL; AF004432; G305016; -.
DR PROSITE; PS00260; GLUCAGON; 5.
DR PFAM; PF00123; hormone2; 5.
SQ SEQUENCE 266 AA; 30951 MW; E6139A25 CRC32;

Query Match 63.28; Score 146; DB 13; Length 266;
Best Local Similarity 60.08; Pred. No. 9.94e-12;
Matches 18; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Db 97 HAEGTFTSDVTQQLDEKAAKEFDWLINGG 126

QY 1 hgegtfslskmqeeavrlfiewlknng 30

RESULT 2
ID Q42144 PRELIMINARY; PRT; 219 AA.
AC Q42144;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PROGLUCAGON II.
OC XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.

```

DR PFAM: PF00123; hormone2; 2.
FT NON_TER 1
SQ SEQUENCE 72 AA; 8293 MW; 0F7AF3EC CRC32;

Query Match 51.9%; Score 120; DB 13; Length 72;
Best Local Similarity 44.8%; Pred. No. 4.25e-07;
Matches 13; Conservative 11; Mismatches 5; Indels 0; Gaps 0

Db 39 HADGTYTSDVSTYLQDQAADFVSWLKS 67
QY 1 hgegtftsdlskqmeeeavrflfkwlng 29
      |||:||||:| :||:| :| :|||:|
      |hgegtftsdlskqmeeeavrflfkwlng 29

RESULT 5
ID Q91408 PRELIMINARY; PRT; 72 AA.
AC Q91408;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PROGLUCAGON (FRAGMENT).
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIRDNERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;
ON ONCORHYNCHUS.
RN [1]
RC SEQUENCE FROM N.A.
RP MEDLINE; 95295739.
RA IRWIN D.M., WONG J.;
RT "Trout and chicken proglucagon: alternative splicing generates mRNA
RT transcripts encoding glucagon-like peptide 2.";
RL MOL. ENDOCRINOL. 9:267-277(1995).
DR EMBL; S78473; G999383; -.
DR PFAM: PF00123; hormone2; 2.
FT NON_TER 1
FT NON_TER 72
SQ SEQUENCE 72 AA; 8293 MW; 0F7AF3EC CRC32;

Query Match 51.9%; Score 120; DB 13; Length 72;
Best Local Similarity 44.8%; Pred. No. 4.25e-07;
Matches 13; Conservative 11; Mismatches 5; Indels 0; Gaps 0

Db 39 HADGTYTSDVSTYLQDQAADFVSWLKS 67
QY 1 hgegtftsdlskqmeeeavrflfkwlng 29
      |||:||||:| :||:| :| :|||:|
      |hgegtftsdlskqmeeeavrflfkwlng 29

RESULT 6
ID Q91971 PRELIMINARY; PRT; 178 AA.
AC Q91971;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE GLUCAGON I..
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIRDNERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;
ON ONCORHYNCHUS.
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE-INTESTINE, DISTAL PORTION;
RC MEDLINE; 95295739.
RA IRWIN D.M., WONG J.;
RT "Trout and chicken proglucagon: alternative splicing generates mRNA
RT transcripts encoding glucagon-like peptide 2.";
RL MOL. ENDOCRINOL. 9:267-277(1995).
DR EMBL; U19917; G736365; -.
DR EMBL; S78475; G999385; -.
DR PROSITE; PS00260; GLUCAGON; 3.
DR PFAM: PF00123; hormone2; 3.
SQ SEQUENCE 178 AA; 20034 MW; 2056F963 CRC32;

Query Match 51.9%; Score 120; DB 13; Length 178;
Best Local Similarity 44.8%; Pred. No. 4.25e-07;

```

DI
UL-500K-139% (KREMERL 06); LASTI (PRODUCTION OF DRIEL)
DE L. PROTEIN, RNA DEPENDENT RNA POLYMERASE.
OS RABIES VIRUS.
GN VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; MONONEGAVIRALES; RHADDOVIRIDAE;
ON LYSSAVIRUS.
OC [1]
RN SEQUENCE FROM N.A.
RP STRAIN-RC-HL;
RA MINAMOTO N.:
RA

RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
 RI "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RI elegans.";
 RI NATURE 368:32-38(1994).
 DR EMBL: 248795; E1351719; -;
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 SQ SEQUENCE 414 AA; 47188 MW; 73B4DF09 CRC32;

Query Match 33.3%; Score 77; DB 5; Length 414;
 Best Local Similarity 35.7%; Pred. No. 2.85e+00;
 Matches 10; Conservative 9; Mismatches 8; Indels 1; Gaps 1;

QY 33 EGDFMYTKKQLDINKLQLFLAVRLKKG 60
 ||| : ||| : ||| : ||| :
 QY 3 egtftsdiskgmeeavrlfie-wlknng 29

Search completed: Mon Oct 4 15:25:30 1999
 Job time : 15 secs.

(ME)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Claim 8
 $X_1 \times X_2 = SD$

geneseq35
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Query		Length	DB	ID	Description	Pred. No.
	Score	Match					
1	260	93.5	39	39	R61769	Exendin-3, for use in	2.62e-15
2	260	93.5	39	14	R80545	Heloderma horridum ex	2.62e-15
3	260	93.5	39	30	W47608	Gila monster extendin-	2.62e-15
4	256	92.1	39	39	R61770	Exendin-4, for use in	6.24e-15
5	256	92.1	39	14	R80545	Heloderma suspectum e	6.24e-15
6	256	92.1	39	30	W47609	Gila monster extendin-	6.24e-15
7	256	92.1	87	35	W70288	Heloderma suspectum p	6.24e-15
8	231	83.1	39	39	R61773	Leu(14), Phe(25)-exen	1.39e-12
9	220	79.1	31	14	R80547	Heloderma suspectum e	1.47e-11
10	209	75.2	31	14	R80543	Heloderma suspectum e	1.54e-10
11	201	72.3	30	39	R61771	Exendin-4 (1-30) for	8.41e-10
12	201	72.3	31	14	R80544	Heloderma suspectum e	8.41e-10
13	198	71.2	30	29	W93901	H. horridum extendin-3	1.59e-09
14	198	71.2	30	39	W93968	H. horridum extendin-3	1.59e-09
15	194	69.8	30	29	W93902	H. horridum extendin-4	3.70e-09
16	189	68.0	30	29	W93909	H. horridum extendin-4	1.06e-08

QY 1 hsdgtfstdskqmeeeavrflfwkngpssgappps 40

RESULT 2
ID R80545 standard; peptide; 39 AA.
AC R80545; 13-14; 17pp; English.
DT 27-FEB-1996 (first entry)
DE Heloderma horridum extendin-3.
KW Extendin-3; diabetes mellitus; hyperglycaemia; insulinotropic peptide.
OS Heloderma horridum.
PN US5424286-A.
PD 13-JUN-1995.
PF 24-MAY-1993; 066480.
PR 24-MAY-1993; US-066480.
PA (ENGJ/) ENG J.
Eng J;
WPI: 95-262627/34.
PT Stimulating/inhibiting insulin release with extendin polypeptide(s) -
for treating diabetes mellitus and preventing hyperglycaemia.
PS Claim 5; Columns 13-14; 17pp; English.
CC R80545 is Heloderma horridum extendin-3. It is an
CC insulinotropic peptide, and can therefore be used in the treatment of
CC diabetes mellitus (types I or II), and for the prevention of
CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
CC and insulin-(in)dependent mechanisms.
SQ Sequence 39 AA;

Query Match 93.5%; Score 260; DB 14; Length 39;
Best Local Similarity 97.5%; Pred. No. 2.62e-15;
Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Db 1 hsdgtfstdskqmeeeavrflfwkngpssgappps 39
QY 1 hsdgtfstdskqmeeeavrflfwkngpssgappps 40

RESULT 3
ID W47608 standard; peptide; 39 AA.
AC W47608;
DT 03-JUL-1998 (first entry)
DE Gila monster extendin-3.
KW Extendin agonist; gastric motility; gastric emptying; treatment;
KW spasm; postprandial dumping syndrome; postprandial hyperglycaemia;
KW type 1 diabetes; impaired glucose tolerance; toxin ingestion;
KW obesity; Gila monster venom; extendin-3.
DE Heloderma horridum.
Key Location/Qualifiers
Modified_site 39 /note= "amidated"
W09805351-A1.
PD 12-FEB-1998.
PF 08-AUG-1997; U14199.
PR 08-AUG-1996; US-694954.
PA (AMYL-) AMYLIN PHARM INC.
PI Beley NRA, Gedulin B, Prickett KS, Young AA;
DR WPI: 98-145351/13.
PT Regulating gastrointestinal motility using extendins or their
PT agonists - for treating spasm, diabetic postprandial hyperglycaemia,
PT impaired glucose tolerance etc., also in diagnostic investigations
PS Claims 20 and 21; Fig 1; 70pp; English.
CC W47549 describes a generic extendin agonist, provided that it does
CC have the formula of either extendin-3 (W47608) or extendin-4
CC (W47609).

CC Extendin agonists, which reduce gastric motility and delay gastric
CC emptying, can be used to treat spasm (where associated with acute
CC diverticulitis or disorders of the biliary tract or sphincter of
CC Oddi), postprandial dumping syndrome and hyperglycaemia
CC (particularly associated with type 2 diabetes), type 1 diabetes,
CC impaired glucose tolerance, toxin ingestion (an extendin agonist is
CC administered to prevent stomach contents passing into the
CC intestines, then the stomach pumped) and obesity. They can also be
CC administered to subjects undergoing gastrointestinal diagnostic
CC investigation, particularly radiological or by magnetic resonance

CC imaging, components of Gila monster venom, have some sequence
CC similarity to glucagon-like peptides (GLP). They are GLP agonists
CC and have been suggested (US5424286) for treatment of diabetes and
CC prevention of hyperglycaemia.
SQ Sequence 39 AA;
Query Match 93.5%; Score 260; DB 30; Length 39;
Best Local Similarity 97.5%; Pred. No. 2.62e-15;
Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Db 1 hsdgtfstdskqmeeeavrflfwkngpssgappps 39
QY 1 hsdgtfstdskqmeeeavrflfwkngpssgappps 40

RESULT 4
ID W61770 standard; peptide; 39 AA.
AC W61770;
DT 29-MAR-1999 (first entry)
DE Extendin-4, for use in treating disorders related to food intake.
KW Extendin; obesity; type II diabetes; eating disorders; cardiac disease;
KW insulin resistance syndrome; elevated plasma glucose level; agonist.
OS Heloderma suspectum.
PN W09830231-A1.
PD 16-JUL-1998.
PF 07-JAN-1998; U00449.
PR 14-NOV-1997; US-066029.
PR 07-JAN-1997; US-034905.
PR 08-AUG-1997; US-055404.
PR 14-NOV-1997; US-065442.
PA (AMYL-) AMYLIN PHARM INC.
PI Beley NRA, Bhavsar S, Prickett KS;
DR WPI: 98-398796/34.
PT Reducing food intake by administering extendins or their
PT analogues - for treatment of e.g. obesity, type II diabetes,
PT eating disorders and insulin resistance
PS Claims 17, 25; Page 8; 214pp; English.
CC The invention relates to a new method for treating disorders that
CC are alleviated by reducing food intake, in particular obesity, type
CC II diabetes, eating disorders, insulin resistance syndrome, elevated
CC plasma glucose levels, or the risk of cardiac disease. The method
CC comprises administering an extendin or an extendin agonist. The treatment
CC reduces appetite and lowers plasma lipid levels. It inhibits food
CC consumption as effectively as amylin or cholecystokinin but has a much
CC longer-lasting action (still effective after 6 hours in a mouse model).
CC The present sequence is that of extendin-4 which is one of the preferred
CC compounds for use in the method.
SQ Sequence 39 AA;

Query Match 92.1%; Score 256; DB 39; Length 39;
Best Local Similarity 92.5%; Pred. No. 6.24e-15;
Matches 37; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
Db 1 hsdgtfstdskqmeeeavrflfwkngpssgappps 39
QY 1 hsdgtfstdskqmeeeavrflfwkngpssgappps 40

RESULT 5
ID R80546 standard; peptide; 39 AA.
AC R80546;
DT 27-FEB-1996 (first entry)
DE Heloderma suspectum extendin-4.
KW Extendin-4; diabetes mellitus; hyperglycaemia; insulinotropic peptide.
OS Heloderma suspectum.
PN US5424286-A.
PD 13-JUN-1995.
PF 24-MAY-1993; 066480.
PR 24-MAY-1993; US-066480.
PA (ENGJ/) ENG J.
PI Eng J;
DR WPI: 95-262627/34.

PT Stimulating/inhibiting insulin release with extendin polypeptide(s) -
 PT for treating diabetes mellitus and preventing hyperglycaemia.
 PS Claim 6; Columns 13-14; 17pp; English.
 CC R80546 is Heloderma suspectum extendin-4. It is an
 CC insulinotropic peptide, and can therefore be used in the treatment of
 CC diabetes mellitus (types I or II), and for the prevention of
 CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
 CC and insulin-(in)dependent mechanisms.
 SQ Sequence 39 AA;

Query Match 92.1%; Score 256; DB 14; Length 39;
 Best Local Similarity 92.5%; Pred. No. 6.24e-15;
 Matches 37; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Db 1 hgegtf-tsldskmqeeavrlfiewlknpgssgappps 39
 QY 1 hsdgtfitsldskmqeeavrlfiewlknpgssgappps 40

RESULT 6

ID W47609 standard; peptide; 39 AA.
 AC W47609;
 DT 03-JUL-1998 (first entry)
 DE Gila monster extendin-4.
 KW Extendin agonist; gastric motility; gastric emptying; treatment;
 KW spasm; postprandial dumping syndrome; postprandial hyperglycaemia;
 KW type 1 diabetes; impaired glucose tolerance; toxin ingestion;
 KW obesity; Gila monster venom; extendin-4.
 OS Heloderma suspectum.
 FH Key Location/Qualifiers
 FT Modified_site 39 /note= "amidated"

PN WO9805351-A1.
 PD 12-FEB-1998.
 PF 08-AUG-1997; U14199.
 PR 08-AUG-1996; US-694954.
 PA (AMYL-) AMYLIN PHARM INC.
 PI Bealey NRA, Gedulin B, Prickett KS, Young AA;
 PT WPI: 98-145351/13.
 DR Regulating gastrointestinal motility using extendins or their
 PT agonists - for treating spasm, diabetic postprandial hyperglycaemia,
 PT impaired glucose tolerance etc., also in diagnostic investigations
 CC Claims 20 and 21; Fig 1; 70pp; English.
 CC W47549 describes a generic extendin agonist, provided that it does
 CC have the formula of either extendin-3 (W47608) or extendin-4
 CC (W47609).
 CC Extendin agonists, which reduce gastric motility and delay gastric
 CC emptying, can be used to treat spasm (where associated with acute
 CC diverticulitis or disorders of the biliary tract or sphincter of
 CC Oddi), postprandial dumping syndrome and hyperglycaemia
 CC (particularly associated with type 2 diabetes), type 1 diabetes,
 CC impaired glucose tolerance, toxin ingestion (an extendin agonist is
 CC administered to prevent stomach contents passing into the
 CC intestines, then the stomach pumped) and obesity. They can also be
 CC administered to subjects undergoing gastrointestinal diagnostic
 CC investigation, particularly radiological or by magnetic resonance
 CC imaging.
 CC Extendins, components of Gila monster venom, have some sequence
 CC similarity to glucagon-like peptides (GLP). They are GLP agonists
 CC and have been suggested (US5424286) for treatment of diabetes and
 CC prevention of hyperglycaemia.
 SQ Sequence 39 AA;

Query Match 92.1%; Score 256; DB 30; Length 39;
 Best Local Similarity 92.5%; Pred. No. 6.24e-15;
 Matches 37; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Db 1 hgegtf-tsldskmqeeavrlfiewlknpgssgappps 39
 QY 1 hsdgtfitsldskmqeeavrlfiewlknpgssgappps 40

RESULT 7

ID W70288 standard; Protein; 87 AA.
 AC W70288;
 DT 06-NOV-1998 (first entry)
 DE Heloderma suspectum proextendin peptide.
 KW Heloderma suspectum proextendin; extendin N-terminal peptide; ENTP;
 KW extendin 4 peptide; extendin 3 gene; Heloderma horridum; metabolic disease;
 KW drug screening; endocrine tumour; organ failure; cell metabolism;
 KW diabetes; reptilian venom peptide.
 OS Heloderma suspectum.
 FH Key Location/Qualifiers
 FT Peptide 1..23 /note= "Signal peptide"
 FT Peptide 1..47 /note= "ENTP"
 FT Peptide 48..87 /note= "Extendin 4"
 FT Cleavage_site 46..47 /note= "Dipeptidyl peptidase cleavage site"

PN WO9835033-A1.
 PD 13-AUG-1998.
 PF 04-FEB-1998; CA0071.
 PR 07-FEB-1997; GB-002582.
 PR 05-FEB-1997; US-037412.
 PA (ONEO-) 1149336 ONTARIO INC.
 PI Drucker DJ;
 DR WPI: 98-447230/38.
 DR N-PSDB; V33163.
 PT New nucleic acid encoding proextendin - used to diagnose and treat,
 PT e.g. endocrine tumours, also to treat poisoning by reptile venom
 PS Claim 3; Fig 2; 26pp; English.
 CC The Heloderma suspectum proextendin peptide is encoded by its cDNA
 CC which was isolated from a H. suspectum salivary gland cDNA library.
 CC The proextendin protein comprises of a novel extendin N-terminal
 CC peptide (ENTP) linked to the N-terminus of the extendin 4 peptide
 CC by a consensus dipeptidyl peptidase cleavage site. The proextendin
 CC cDNA can be used to clone or identify related sequences (e.g. the
 CC extendin 3 gene of Heloderma horridum, mutant alleles and proextendin
 CC gene regulatory defects associated with metabolic disease) and species
 CC homologues (e.g. for developing animal models for drug screening).
 CC The proextendin peptide can be used to raise antibodies. Anti-proextendin
 CC antibodies are claimed to be useful for diagnosing conditions associated
 CC with altered levels of proextendin (e.g. endocrine tumours and organ
 CC failure), for identifying other regulators of cell metabolism, in drug
 CC screens and for treating metabolic diseases (e.g. diabetes) and for
 CC neutralising, or detecting, reptilian venom peptides.
 SQ Sequence 87 AA;

Query Match 92.1%; Score 256; DB 35; Length 87;
 Best Local Similarity 92.5%; Pred. No. 6.24e-15;
 Matches 37; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Db 48 hgegtf-tsldskmqeeavrlfiewlknpgssgappps 86
 QY 1 hsdgtfitsldskmqeeavrlfiewlknpgssgappps 40

RESULT 8

ID W61773 standard; peptide; 39 AA.
 AC W61773;
 DT 29-MAR-1999 (first entry)
 DE Leu(14), Phe(25)-extendin-4 amide, for reducing food intake.
 KW Extendin; obesity; type II diabetes; eating disorders; cardiac disease;
 KW insulin resistance syndrome; elevated plasma glucose level; agonist.
 OS Heloderma suspectum.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Modified_site 39 /note= "the C-terminal is in amide form"

PN WO9830231-A1.
 PD 16-JUL-1998.
 PF 07-JAN-1998; U00449.
 PR 14-NOV-1997; US-066029.
 PR 07-JAN-1997; US-034905.


```

PR 08-AUG-1997; US-055404.
PR 14-NOV-1997; US-065442.
PA (AMYL-) AMYLIN PHARM INC.
PI Bealey NRA, Bhavsar S, Prickett KS;
DR WPI: 98-398796/34.
PT Reducing food intake by administering extendins or their
PT analogues - for treatment of e.g. obesity, type II diabetes,
PT eating disorders and insulin resistance
PS Claims 18, 26; Page 12; 214pp; English.
CC The invention relates to a new method for treating disorders that
CC are alleviated by reducing food intake, in particular obesity, type
CC II diabetes, eating disorders, insulin resistance syndrome, elevated
CC plasma glucose levels, or the risk of cardiac disease. The method
CC comprises administering an extendin or an extendin agonist. The treatment
CC reduces appetite and lowers plasma lipid levels. It inhibits food
CC consumption as effectively as amylin or cholecystokinin but has a much
CC longer-lasting action (still effective after 6 hours in a mouse model).
CC The present sequence is that of an extendin-4 variant which is one of
CC the preferred compounds for use in the method.
SQ Sequence 39 AA;

Query Match 83.1%; Score 231; DB 39; Length 39;
Best Local Similarity 87.5%; Pred. No. 1.39e-12;
Matches 35; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

Db 1 hgegtf-tsdlskqleeeavrfliefknpgssgappgs 39
QY 1 hsdgtfitsdkqmeeeavrfliefknpgssgappgs 40

RESULT 9
ID R80547;
AC R80547;
DT 27-FEB-1996 (first entry)
DE Heloderma suspectum extendin-4 residues 9-39 (Extendin-4(9-39)).
KW Extendin-4; Residues 9-39; Extendin-4(9-39);
KW Insulinotropic peptides; inhibitor.
OS Heloderma suspectum.
PN US5424286-A.
PD 13-JUN-1995.
PF 24-MAY-1993; 066480.
PR 24-MAY-1993; US-066480.
PA (ENGJ/) ENG J.
PI Eng J;
PT WPI: 95-262627/34.
PT Stimulating/inhibiting insulin release with extendin polypeptide(s) -
PT for treating diabetes mellitus and preventing hyperglycaemia.
PS Claim 7; Columns 13-14; 17pp; English.
CC R80547 is the Heloderma suspectum extendin-4 residues 9-39. It
CC is an insulinotropic peptide activity inhibitor.
SQ Sequence 31 AA;

Query Match 79.1%; Score 220; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.47e-11;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 dlskqmeeeavrfliefknpgssgappgs 31
QY 10 dlskqmeeeavrfliefknpgssgappgs 40

RESULT 10
ID R80543;
AC R80543;
DT 27-FEB-1996 (first entry)
DE Heloderma suspectum extendin-4 residues 1-31 (Extendin-4(1-31)).
KW Extendin-4; Residues 1-31; Extendin-4(1-31); diabetes mellitus;
KW hyperglycaemia; insulinotropic peptide.
OS Heloderma suspectum.
PN US5424286-A.
PD 13-JUN-1995.
PF 24-MAY-1993; 066480.
PR 24-MAY-1993; US-066480.
PA (ENGJ/) ENG J.
PI Eng J;
PT WPI: 95-262627/34.
PT Stimulating/inhibiting insulin release with extendin polypeptide(s) -
PT for treating diabetes mellitus and preventing hyperglycaemia.
PS Claim 1; Columns 13-14; 17pp; English.
CC R80543 is the Heloderma suspectum extendin-4 residues 1-31. It is an
CC insulinotropic peptide, and can therefore be used in the treatment of
CC diabetes mellitus (types I or II), and for the prevention of
CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
CC and insulin-(in)dependent mechanisms.
SQ Sequence 31 AA;

Query Match 75.2%; Score 209; DB 14; Length 31;
Best Local Similarity 90.6%; Pred. No. 1.54e-10;
Matches 29; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Db 1 hgegtf-tsdlskqmeeeavrfliefknpgssg 31
QY 1 hsdgtfitsdkqmeeeavrfliefknpgssg 32

RESULT 11
ID W61771 standard; peptide; 30 AA.
AC W61771;
DT 29-MAR-1999 (first entry)
DE Extendin-4 (1-30) for use in treating disorders related to food intake.
KW Extendin; obesity; type II diabetes; eating disorders; cardiac disease;
KW insulin resistance syndrome; elevated plasma glucose level; agonist.
OS Heloderma suspectum.
FH Key
FT Location/Qualifiers
FT Modified_site 30
FT FT /note= "Optionally the C-terminal is in amide form"
PN WQ9830231-A1.
PD 16-JUL-1998.
PF 07-JAN-1998; U00449.
PR 14-NOV-1997; US-066029.
PR 07-JAN-1997; US-034905.
PR 08-AUG-1997; US-055404.
PR 14-NOV-1997; US-065442.
PA (AMYL-) AMYLIN PHARM INC.
PI Bealey NRA, Bhavsar S, Prickett KS;
DR WPI: 98-398796/34.
PT Reducing food intake by administering extendins or their
PT analogues - for treatment of e.g. obesity, type II diabetes,
PT eating disorders and insulin resistance
PS Claims 18, 26; Page 11; 214pp; English.
CC The invention relates to a new method for treating disorders that
CC are alleviated by reducing food intake, in particular obesity, type
CC II diabetes, eating disorders, insulin resistance syndrome, elevated
CC plasma glucose levels, or the risk of cardiac disease. The method
CC comprises administering an extendin or an extendin agonist. The treatment
CC reduces appetite and lowers plasma lipid levels. It inhibits food
CC consumption as effectively as amylin or cholecystokinin but has a much
CC longer-lasting action (still effective after 6 hours in a mouse model).
CC The present sequence is that of extendin-4 (1-30) or its amide which is
CC one of the preferred compounds for use in the method.
SQ Sequence 30 AA;

Query Match 72.3%; Score 201; DB 39; Length 30;
Best Local Similarity 90.3%; Pred. No. 8.41e-10;
Matches 28; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Db 1 hgegtf-tsdlskqmeeeavrfliefknpgss 30
QY 1 hsdgtfitsdkqmeeeavrfliefknpgss 31

RESULT 12
ID R80544 standard; peptide; 31 AA.
AC R80544;
DT 27-FEB-1996 (first entry)
DE Heloderma suspectum extendin-4 residues 1-31-Tyr31.

```

KW Exendin-4; residues 1-31; Y-31-Exendin-4(1-31); diabetes mellitus;
 KW hyperglycaemia; Tyr31; insulinotropic peptide.
 OS Heloderma suspectum.
 PN US5424286-A.
 IN 13-JUN-1995.
 PR 24-MAY-1993; 066480.
 PA (ENGJ/) ENG J.
 PI Eng J;
 DR WPI; 95-262627/34.
 PT Stimulating/inhibiting insulin release with exendin polypeptide(s) -
 PT for treating diabetes mellitus and preventing hyperglycaemia.
 PT Claim 2; Columns 13-14; 17pp; English.
 CC R00544 is the Heloderma suspectum exendin-4 residues 1-31, where
 CC the native Pro31 has been replaced with a Tyr residue. It is an
 CC insulinotropic peptide, and can therefore be used in the treatment of
 CC diabetes mellitus (types I or II), and for the prevention of
 CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
 CC and insulin-(in)dependent mechanisms.
 CC Sequence 31 AA;

Query Match 72.3%; Score 201; DB 14; Length 31;
 Best Local Similarity 90.3%; Pred. No. 8,41e-10;
 Matches 28; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Db 1 hsgdtf-tsldskmqeeavrlfiewlknng 30
 QY 1 hsdgtf-tsldskmqeeavrlfiewlknng 31

RESULT 13

ID W39301 standard; peptide; 30 AA.
 AC W39301;
 DT 05-JUN-1998 (first entry)
 DE H. horridum exendin-3 peptide.
 KW Exendin-3; exendin 4; diabetes; insulin; secretion; biosynthesis;
 KW glucagon reduction; hypoglycaemia; glucose; treatment.
 OS Heloderma horridum.
 FH Key Location/Qualifiers
 FT Modified_site 30
 FT /note= "This residue can be any amino acid except Gly"
 PN W09746584-A1.
 PD 11-DEC-1997.
 PF 03-JUN-1997; E02930.
 PR 13-SEP-1996; DE-037230.
 PR 05-JUN-1996; DE-022502.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Goeke B, Goeke R, Hoffmann E;
 DR WPI; 98-042119/04.
 CC Truncated versions of exendin peptide(s) for treating diabetes -
 CC increase secretion and biosynthesis of insulin, but reduce those of
 CC glucagon, and do not induce hypoglycaemia
 CC Claim 1; Page 3; 150pp; English.
 CC This peptide is a fragment of exendin-3 isolated from Heloderma
 CC horridum. This peptide and its salts, esters and derivatives can be
 CC used to treat diabetes mellitus. They stimulate biosynthesis and
 CC secretion of insulin, but have the opposite effect on glucagon, and
 CC independent of this activity can increase peripheral glucose utilisation.
 CC Exendin-3 and exendin-4 are only active when blood sugar levels are
 CC high, so they will not induce hypoglycaemia. Compared with glucagon-like
 CC peptide 1 (GLP1) and the known exendins, they are more active (effective
 CC at lower doses), more stable to degradation and metabolism and have a
 CC longer lasting effect. Truncated forms of this peptide can be made more
 CC economically than full length versions.
 CC Sequence 30 AA;

Query Match 71.2%; Score 198; DB 29; Length 30;
 Best Local Similarity 96.7%; Pred. No. 1.59e-09;
 Matches 29; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 hsdgtf-tsldskmqeeavrlfiewlknng 29
 QY 1 hsdgtf-tsldskmqeeavrlfiewlknng 30

QY 1 hsdgtf-tsldskmqeeavrlfiewlknng 30
 RESULT 14
 ID W39368 standard; peptide; 30 AA.
 AC W39368;
 DT 05-JUN-1998 (first entry)
 DE H. horridum exendin-3 peptide derivative #11.
 KW Exendin-3; exendin 4; diabetes; insulin; secretion; biosynthesis;
 KW glucagon reduction; hypoglycaemia; glucose; treatment.
 OS Heloderma horridum.
 FH Key Location/Qualifiers
 FT Modified_site 30
 FT /note= "C-terminal amide"

PN W09746584-A1.
 PD 11-DEC-1997.
 PF 03-JUN-1997; E02930.
 PR 13-SEP-1996; DE-037230.
 PR 05-JUN-1996; DE-022502.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Goeke B, Goeke R, Hoffmann E;
 DR WPI; 98-042119/04.
 CC Truncated versions of exendin peptide(s) for treating diabetes -
 CC increase secretion and biosynthesis of insulin, but reduce those of
 CC glucagon, and do not induce hypoglycaemia
 CC Claim 2; Page 27; 150pp; English.
 CC Peptides W39303-W39420 are fragments of exendin-3 and exendin-4
 CC isolated from Heloderma horridum which are used in a novel method
 CC for the treatment of diabetes mellitus. These peptides can stimulate
 CC biosynthesis and secretion of insulin, but have the opposite effect on
 CC glucagon, and independent of this activity can increase peripheral
 CC sugar levels are high, so they will not induce hypoglycaemia. Compared
 CC with glucagon-like peptide 1 (GLP1) and the known exendins, they are more
 CC active (effective at lower doses), more stable to degradation and
 CC metabolism and have a longer lasting effect. Truncated forms of this
 CC peptide can be made more economically than full length versions.
 CC Sequence 30 AA;

Query Match 71.2%; Score 198; DB 29; Length 30;
 Best Local Similarity 96.7%; Pred. No. 1.59e-09;
 Matches 29; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 hsdgtf-tsldskmqeeavrlfiewlknng 29
 QY 1 hsdgtf-tsldskmqeeavrlfiewlknng 30

RESULT 15

ID W39302 standard; peptide; 30 AA.
 AC W39302;
 DT 05-JUN-1998 (first entry)
 DE H. horridum exendin-4 peptide.
 KW Exendin-3; exendin 4; diabetes; insulin; secretion; biosynthesis;
 KW glucagon reduction; hypoglycaemia; glucose; treatment.
 OS Heloderma horridum.
 FH Key Location/Qualifiers
 FT Modified_site 30
 FT /note= "This residue can be any amino acid except Gly"
 PN W09746584-A1.
 PD 11-DEC-1997.
 PF 03-JUN-1997; E02930.
 PR 13-SEP-1996; DE-037230.
 PR 05-JUN-1996; DE-022502.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Goeke B, Goeke R, Hoffmann E;
 DR WPI; 98-042119/04.
 CC Truncated versions of exendin peptide(s) for treating diabetes -
 CC increase secretion and biosynthesis of insulin, but reduce those of
 CC glucagon, and do not induce hypoglycaemia
 CC Claim 1; Page 4; 150pp; English.
 CC This peptide is a fragment of exendin-4 isolated from Heloderma

CC horridum. This peptide and its salts, esters and derivatives can be
CC used to treat diabetes mellitus. They stimulate biosynthesis and
CC secretion of insulin, but have the opposite effect on glucagon, and
CC independent of this activity can increase peripheral glucose utilisation.
CC Exendin-3 and exendin-4 are only active when blood sugar levels are
CC high, so they will not induce hypoglycaemia. Compared with glucagon-like
CC peptide 1 (GLP1) and the known exendins, they are more active (effective
CC at lower doses), more stable to degradation and metabolism and have a
CC longer lasting effect. Truncated forms of this peptide can be made more
CC economically than full length versions.
SQ Sequence : 30 AA;

Query Match 59.8%; Score 194; DB 29; Length 30;
Best Local Similarity 90.0%; Pred. No. 3.70e-09;
Matches 27; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

1 hgegtf-tsdlskqmeeeavrlfiewlkng 29
::||| ||||| ||||| ||||| |||||
47 1 hsdgtfitsdlskqmeeeavrlfiewlkng 30

Search completed: Mon Oct 4 15:30:39 1999
Job time : 20 secs.

MPQRH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Oct 4 15:29:52 1999; MasPar time 4.88 Seconds
328.404 Million cell updates/sec
Abular output not generated.

le: >MOHAM-312-CLAIM83A.PEP
Description: (1-40) from moham312177.ppe
Sequence: 1 hsdgtfitsdlskqmeeeavrlfiewlknpgssgappps 40

Scoring table: PAM 150
Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir60
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 35.024; Variance 64.745; scale 0.541

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description	Pred. No.
1	250	93.5	39	1	HWGH32	exendin-3 - Mexican b	7.55e-34		
2	256	92.1	39	1	HWGH4G	exendin-4 - Gila mons	4.98e-33		
3	118	42.4	101	1	GCFCB	glucagon precursor -	1.45e-06		
4	117	42.1	31	2	S44472	glucagon G2 - North A	2.15e-06		
5	116	41.7	63	1	GCIDC	glucagon precursor -	3.18e-06		
6	113	40.6	30	2	S44473	glucagon-like peptide	1.02e-05		
7	112	40.3	31	2	S44471	glucagon G1 - North A	1.49e-05		
8	111	39.9	66	2	I51093	glucagon - chinook sa	2.19e-05		
9	111	39.9	178	2	I51058	glucagon I precursor	2.19e-05		
10	111	39.9	178	2	I51057	glucagon II precursor	2.19e-05		
11	110	39.6	30	2	C61125	glucagon-like peptide	3.21e-05		
12	110	39.6	30	2	B61125	glucagon-like peptide	3.21e-05		
13	110	39.6	72	1	GCXGA	glucagon precursor -	3.21e-05		
14	109	39.2	122	1	GCXFA	glucagon 2 precursor	4.70e-05		
15	108	38.8	60	1	GCXNC	glucagon precursor -	6.87e-05		
16	106	38.1	29	1	GCDF	glucagon - smaller sp	1.46e-04		
17	104	37.4	124	1	GCAF	glucagon I precursor	3.09e-04		
18	104	37.4	158	1	GCGP	glucagon precursor -	3.09e-04		
19	104	37.4	180	1	GCRT	glucagon precursor -	3.09e-04		
20	104	37.4	180	1	GCRDTU	glucagon precursor -	3.09e-04		
21	104	37.4	180	1	GCBO	glucagon precursor -	3.09e-04		
22	104	37.4	180	1	GCHY	glucagon precursor -	3.09e-04		
23	104	37.4	180	1	GCGP	glucagon precursor -	3.09e-04		

24	104	37.4	180	1	GCHU	glucagon precursor -	3.09e-04
25	104	37.4	150	2	A57294	glucagon precursor -	3.09e-04
26	103	37.1	151	1	GCCH	glucagon precursor -	4.47e-04
27	103	37.1	206	2	I51301	proglucagon - chicken	4.47e-04
28	100	36.0	29	1	GCCB	glucagon - Chinchilla	1.35e-03
29	99	35.6	29	2	S07211	glucagon - marbled el	1.95e-03
30	95	34.2	29	1	GCFLS	glucagon - European f	8.25e-03
31	95	34.2	29	2	A61135	glucagon - bigeye tun	8.25e-03
32	95	34.2	87	1	GCFLS	glucagon precursor -	8.25e-03
33	90	32.4	29	2	A91741	glucagon - rabbit (te	4.81e-02
34	90	32.4	29	2	C39258	glucagon - common squ	4.81e-02
35	90	32.4	29	1	GCEN	glucagon - elephanti	4.81e-02
36	90	32.4	29	2	A91742	glucagon - Arabian ca	4.81e-02
37	90	32.4	69	1	GCDF69	glucagon-69 - dog	4.81e-02
38	89	32.0	29	2	S39018	glucagon - bowfin	6.80e-02
39	88	31.7	29	2	C60840	glucagon I - European	9.59e-02
40	88	31.7	2127	1	ZLVNSB	genome polyprotein -	9.59e-02
41	88	31.7	2142	1	ZLVNVP	genome polyprotein -	9.59e-02
42	86	30.9	29	2	A91740	glucagon - turkey (te	1.90e-01
43	86	30.9	29	1	GCPVP	glucagon - North Amer	1.90e-01
44	86	30.9	552	2	S45978	replicase - phage PP7	1.90e-01
45	85	30.6	29	1	A61583	glucagon - ostrich	2.66e-01

ALIGNMENTS

RESULT 1
ENTRY
TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#journal
#title
#cross-references MWID:91056067
#accession A23674
#molecule_type protein
#residues 1-39 ##label ENG
COMMENT Exendins are venom components that are thought to bind to receptors
for vasoactive intestinal peptide and/or secretin on pancreatic
acinar cells and to activate adenylate cyclase, resulting in
secretion of amylase.
CLASSIFICATION #superfamily glucagon
KEYWORDS amidated carboxyl end; duplication; secretagogue; venom
FEATURE
39
SUMMARY #length 39 #molecular-weight 4204 #checksum 9591
Query Match 93.5%; Score 260; DB 1; Length 39;
Best Local Similarity 97.5%; Pred. No. 7.55e-34;
Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Db 1 HSDGTF-TSDLSKQMEEEAVRLFIEWLKNPGSSGAPPPS 39
QY 1 hsdgtfitsdlskqmeeeavrlfiewlknpgssgappps 40

RESULT 2
ENTRY
TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
HWGH4G #type complete
exendin-4 - Gila monster
#formal_name Heloderma suspectum #common_name Gila monster
31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
21-Nov-1997
A42486
A42486
Eng, J.; Kleinman, W.A.; Singh, L.; Singh, G.; Raufman, J.P.


```

ENTRY           S44473      #type complete
TITLE          glucagon-like peptide - North American paddlefish (Polyodon
ORGANISM       spathula)
DATE           18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change
              20-Mar-1998
ACCESSIONS     S44473
REFERENCE      S44467
#authors      Nguyen, T.M.; Mommensen, T.P.; Mims, S.M.; Conlon, J.M.
#journal      Biochem. J. (1994) 300:339-345
#title        Characterization of insulins and proglucagon-derived peptides
              from a phylogenetically ancient fish, the paddlefish
              (Polyodon spathula).
#accession    S44473
#molecule_type protein
#residues     1-30 #label NGU
CLASSIFICATION #superfamily glucagon
KEYWORDS       duplication; hormone; pancreas
FEATURE        1-30
               #product glucagon-like peptide #status predicted #label
SUMMARY         MAT
               #length 30 #molecular-weight 3359 #checksum 5186
Query Match   40.6%; Score 113; DB 2; Length 30;
Best Local Similarity 56.7%; Pred. No. 1.02e-05;
Matches 17; Conservative 6; Mismatches 6; Indels 1; Gaps 1;
Db            1 HADGTY-TSDASSFLQEQAAARDFTISWLKKG 29
               ||||| :|::|| |::|| |::|||
QY            1 hsdgtfisdlskqmeeeavrllfwlknng 30
               ||||| :|::|| |::|| |::|||

RESULT        7
ENTRY         S44471      #type complete
TITLE        glucagon G1 - North American paddlefish (Polyodon spathula)
ORGANISM     polyodon
DATE         18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change
              20-Mar-1998
ACCESSIONS   S44471
REFERENCE    S44467
#authors     Nguyen, T.M.; Mommensen, T.P.; Mims, S.M.; Conlon, J.M.
#journal     Biochem. J. (1994) 300:339-345
#title       Characterization of insulins and proglucagon-derived peptides
              from a phylogenetically ancient fish, the paddlefish
              (Polyodon spathula).
#accession   S44471
#molecule_type protein
#residues    1-31 #label NGU
#experimental_source pancreas
CLASSIFICATION #superfamily glucagon
#words        carbohydrate metabolism; duplication; hormone; pancreas
STRUCTURE     1-31
              #product glucagon G1 #status predicted #label MAT
SUMMARY        length 31 #molecular-weight 3751 #checksum 7808
Query Match   40.3%; Score 112; DB 2; Length 31;
Best Local Similarity 56.7%; Pred. No. 1.49e-05;
Matches 17; Conservative 5; Mismatches 7; Indels 1; Gaps 1;
Db            1 HSQCWF-TNDYSKYLEEKRAKEFVWLNKG 29
               ||||| :|::|| |::|| |::|||
QY            1 hsdgftisdlskqmeeeavrllfwlknng 30
               ||||| :|::|| |::|| |::|||

RESULT        8
ENTRY         I51093      #type fragment
TITLE        glucagon - chinook salmon (fragment)
ORGANISM     Oncomorhynchus tshawytscha #common_name chinook
              salmon
DATE         13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
              21-Nov-1997
ACCESSIONS   I51093
REFERENCE    A55895
```

```

1:||||| :||||| :||||| :||||| :|||||
QY 1 hsdgtfslskqmeeeavrlfiewlkg 30

RESULT 10
ENTRY
TITLE glucagon II precursor - rainbow trout
ORGANISM #formal_name Oncorhynchus mykiss #common_name rainbow trout
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
21-Nov-1997
ACCESSIONS I51057; I51039; I51038
REFERENCE A53895
#authors Irwin, D.M.; Wong, J.
#journal Mol. Endocrinol. (1995) 9:267-277
#title Trout and chicken proglucagon: alternative splicing generates
#title mRNA transcripts encoding glucagon-like peptide 2.
#cross-references MUID:95295739
#accession I51057
##status Preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-178 ##label IRW
##cross-references EMBL:U19914; NID:g736362; PID:g736363
#accession I51039
##status Preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 113-144 ##label IR2
##cross-references EMBL:U19916; NID:g736369; PID:g736372
#accession I51038
##status Preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 113-123 ##label IR3
##cross-references EMBL:U19915; NID:g736368; PID:g736371
GENETICS
#introns 123/2
CLASSIFICATION #superfamily glucagon
KEYWORDS duplication
SUMMARY #length 178 #molecular-weight 19998 #checksum 4464

Query Match 39.9%; Score 111; DB 2; Length 178;
Best Local Similarity 46.7%; Pred. No. 2.19e-05;
Matches 14; Conservative 10; Mismatches 5; Indels 1; Gaps 1;

Db 90 HADGTY-TSDVSYLQDQAKFVSWLKG 118
1:||||| :||||| :||||| :||||| :|||||
1 hsdgtfslskqmeeeavrlfiewlkg 30

RESULT 11
ENTRY
TITLE glucagon-like peptide - European eel
ORGANISM #formal_name Anguilla anguilla #common_name European eel
DATE 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change
21-Nov-1997
ACCESSIONS C61125
REFERENCE A61125
#authors Conlon, J.M.; Andrews, P.C.; Thim, L.; Moon, T.W.
#journal Gen. Comp. Endocrinol. (1991) 82:23-32
#title The primary structure of glucagon-like peptide but not
#title insulin has been conserved between the American eel,
#title Anguilla rostrata and the European eel, Anguilla anguilla.
#cross-references MUID:91340068
#accession C61125
##molecule_type protein
##residues 1-30 ##label CON
CLASSIFICATION #superfamily glucagon
KEYWORDS amidated carboxyl end; duplication
FEATURE 1-30
#product glucagon-like peptide #status experimental
#label GLP\
#modified_site amidated carboxyl end (Arg) #status
predicted
SUMMARY #length 30 #molecular-weight 3376 #checksum 6092

Query Match 39.6%; Score 110; DB 2; Length 30;
Best Local Similarity 43.3%; Pred. No. 3.21e-05;
Matches 13; Conservative 10; Mismatches 6; Indels 1; Gaps 1;

Db 1 HAEQTY-TSDVSYLQDQAKFVSWLKG 29
1:||||| :||||| :||||| :||||| :|||||
1 hsdgtfslskqmeeeavrlfiewlkg 30

RESULT 12
ENTRY
TITLE glucagon-like peptide - American eel
ORGANISM #formal_name Anguilla rostrata #common_name American eel
DATE 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change
21-Nov-1997
ACCESSIONS B61125
REFERENCE A61125
#authors Conlon, J.M.; Andrews, P.C.; Thim, L.; Moon, T.W.
#journal Gen. Comp. Endocrinol. (1991) 82:23-32
#title The primary structure of glucagon-like peptide but not
#title insulin has been conserved between the American eel,
#title Anguilla rostrata and the European eel, Anguilla anguilla.
#cross-references MUID:91340068
#accession B61125
##molecule_type protein
##residues 1-30 ##label CON
CLASSIFICATION #superfamily glucagon
KEYWORDS amidated carboxyl end; duplication
FEATURE 1-30
#product glucagon-like peptide #status experimental
#label GLP\
#modified_site amidated carboxyl end (Arg) #status
predicted
SUMMARY #length 30 #molecular-weight 3376 #checksum 6092

Query Match 39.6%; Score 110; DB 2; Length 30;
Best Local Similarity 43.3%; Pred. No. 3.21e-05;
Matches 13; Conservative 10; Mismatches 6; Indels 1; Gaps 1;

Db 1 HAEQTY-TSDVSYLQDQAKFVSWLKG 29
1:||||| :||||| :||||| :||||| :|||||
1 hsdgtfslskqmeeeavrlfiewlkg 30

RESULT 13
ENTRY
TITLE glucagon precursor - alligator gar (fragment)
ORGANISM #formal_name Lepisosteus spatula #common_name alligator gar
DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
20-Mar-1998
ACCESSIONS S06339; S06871
REFERENCE S06339
#authors Pollock, H.G.; Kimmel, J.R.; Ebner, K.E.; Hamilton, J.W.;
#journal Rouse, J.B.; Lance, V.; Rawitch, A.B.
#journal Gen. Comp. Endocrinol. (1988) 69:133-140
#title Isolation of alligator gar (Lepisosteus spatula) glucagon,
#title oxyntomodulin, and glucagon-like peptide: amino acid
#title sequences of oxyntomodulin and glucagon-like peptide.
#cross-references MUID:88196798
#accession S06339
##molecule_type protein
##residues 1-36 ##label POL
#accession S06871
##molecule_type protein
##residues 39-72 ##label P02
COMMENT X's at positions 37-38 represent a pair of basic amino acids
forming a cleavage site.
CLASSIFICATION #superfamily glucagon
KEYWORDS carbohydrate metabolism; duplication; hormone; pancreas
FEATURE

```

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Run on: Mon Oct 4 15:28:55 1999; MasPar time 3.27 Seconds
*****bular output not generated. 346.176 Million cell updates/sec

File: >MOHAM-312-CLAIM83A.PEP
Description: (1-40) from moham312177.ppe
Sequence: 1 hsdgtfstdskqmeeeavrlfiewkngpssgappps 40

Scoring table: PAM 150
Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot37
i:swissprot

Statistics: Mean 35.894; Variance 59.375; scale 0.605

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	260	93.5	39	1	EXE3_HELHO	7.01e-38
2	256	92.1	87	1	EXE4_HELHU	5.73e-37
3	118	42.4	103	1	GLUC_RANCA	1.29e-07
4	117	42.1	71	1	GLUC_ICTPU	1.99e-07
5	110	39.6	30	1	GLUC_ANGAN	3.91e-06
6	110	39.6	78	1	GLUC_LEPSP	3.91e-06
7	109	39.2	122	1	GLUC_LOPAM	5.95e-06
8	108	38.8	68	1	GLUC_ONCKI	9.03e-06
9	108	38.8	121	1	GLUC_CARAU	9.03e-06
10	106	38.1	29	1	GLUC_SCYCA	2.07e-05
11	104	37.4	124	1	GLUC_LOPAM	4.72e-05
12	104	37.4	158	1	GLUC_PIG	4.72e-05
13	104	37.4	180	1	GLUC_BOVIN	4.72e-05
14	104	37.4	180	1	GLUC_MESAU	4.72e-05
15	104	37.4	180	1	GLUC_MESAU	4.72e-05
16	104	37.4	180	1	GLUC_OCTDE	4.72e-05
17	104	37.4	180	1	GLUC_HUMAN	4.72e-05
18	104	37.4	180	1	GLUC_CAVPO	4.72e-05
19	104	37.4	180	1	GLUC_RAT	4.72e-05
20	103	37.1	151	1	GLUC_CHICK	7.11e-05
21	100	36.0	29	1	GLUC_CHIRB	2.40e-04
22	99	35.6	29	1	GLUC_TORMA	3.58e-04
23	96	34.5	96	1	GLUC_MYOSC	1.18e-03

24	95	34.2	29	1	GLUC_PLAPE	1.75e-03
25	94	33.8	33	1	GLUC_ORENI	2.58e-03
26	94	33.8	36	1	GLUC_URENI	2.58e-03
27	90	32.4	29	1	GLUC_CALMI	1.21e-02
28	90	32.4	29	1	GLUC_RABIT	1.21e-02
29	90	32.4	69	1	GLUC_CANFA	1.21e-02
30	89	32.0	75	1	GLUC_AMICA	1.21e-02
31	88	31.7	2127	1	REPL_RABVS	1.77e-02
32	88	31.7	2142	1	REPL_RABVP	2.57e-02
33	86	30.9	29	1	GLUC_DIDMA	5.43e-02
34	85	30.6	29	1	GLUC_ANAPL	7.86e-02
35	84	30.2	36	1	GLUC_HYDGO	1.13e-01
36	84	30.2	406	1	YBDN_ECOLI	1.13e-01
37	80	28.8	290	1	BET4_YEAST	4.79e-01
38	76	27.3	429	1	IFS_RAT	1.93e+00
39	76	27.3	431	1	IFS_HUMAN	1.93e+00
40	76	27.3	611	1	TAP_CHICK	1.93e+00
41	76	27.3	658	1	UVRB_HELPY	1.93e+00
42	76	27.3	1075	1	Y124_METJA	1.93e+00
43	75	27.0	153	1	GIP_HUMAN	2.71e+00
44	75	27.0	389	1	UTL_HUMAN	2.71e+00
45	75	27.0	600	1	TAPL_MOUSE	2.71e+00

ALIGNMENTS

RESULT 1	STANDARD;	PRT;	39 AA.
ID EXE3_HELHO			
AC P20394;			
DT 01-FEB-1991 (REL. 17, CREATED)			
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)			
DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)			
DE EXENDIN-3.			
OS HELODERMA HORRIDUM (MEXICAN BEADED LIZARD).			
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;			
OC SCLEROGLOSSA; ANGIOMORPHA; HELODERMATIDAE; HELODERMA.			
RN [1]			
RP SEQUENCE.			
RC TISSUE=VENOM;			
RX MEDLINE: 91056067.			
RA ENG J., ANDREW P.C., KLEINMAN W.A., SINGH L., RAUFMAN J.-P.;			
RT "Purification and structure of exendin-3, a new pancreatic			
RL J. BIOL. CHEM. 265:20259-20262(1990).			
CC -1- FUNCTION: HAS A VIP/SECRETIN-LIKE BIOLOGICAL ACTIVITY. INTERACTS			
CC WITH THE EXENDIN RECEPTOR.			
CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.			
DR PIR; A23674; HWGH3Z.			
DR PROSITE; PS00260; GLUCAGON; 1.			
DR PFAM; PF00123; hormone2; 1.			
DR HSSP; P01274; IGCN.			
KW GLUCAGON FAMILY; VENOM; AMIDATION.			
FT MOD_RES 39 39			
SQ SEQUENCE 39 AA; 4204 MW; AB598FD3 CRC32;			
Query Match 93.5%; Score 260; DB 1; Length 39;			
Best Local Similarity 97.5%; Pred. No. 7.01e-38;			
Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;			

Db 1	HSNGTPTDLSKQMEEEAVRLFIEMKNGPSSGAPPPS 39		
QY 1	hsdgtfstdskqmeeeavrlfiewkngpssgappps 40		
RESULT 2	STANDARD;	PRT;	87 AA.
ID EXE4_HELHU			
AC P26349;			
DT 01-MAY-1992 (REL. 22, CREATED)			
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)			
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
DE EXENDIN-4 PRECURSOR.			
OS HELODERMA SUSPECTUM (GIILA MONSTER).			
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;			

OC SCLEROGLOSSA; ANGUIMORPHA; HELODERMATIDAE; HELODERMA.
 RN [1]
 CC SEQUENCE FROM N.A.
 CC CHEN Y.E., DRUCKER D.J.:
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 CC SEQUENCE OF 48-86.
 RP TISSUE=VENOM;
 RC MEDLINE; 92218391.
 RX ENG J., KLINKAN W.A., SINGH L., SINGH G., RAUFMAN J.P.:
 RA "Isolation and characterization of extendin-4, an extendin-3 analogue,
 RT from Heloderma suspectum venom. Further evidence for an extendin
 RT receptor on dispersed acini from guinea pig pancreas.";
 RL J. BIOL. CHEM. 267:7402-7405(1992).
 CC -!- FUNCTION: HAS A VIP/SECRETIN-LIKE BIOLOGICAL ACTIVITY. INTERACTS
 CC WITH THE EXTENDIN RECEPTOR.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U77613; G1916067; -.
 DR PIR; A42486; HWHG4G.
 DR PROSITE; PS00250; GLUCAGON; 1.
 DR PFAM; PF00123; hormone2; 1.
 KW GLUCAGON FAMILY; VENOM; AMIDATION; SIGNAL.
 FT SIGNAL 1 23 POTENTIAL.
 FT PEPTIDE 43 86 EXTENDIN-4.
 FT MOD_RES 86 86 AMIDATION (G-87 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 87 AA; 9479 MW; 6C1A9FD5 CRC32;
 CC
 CC Query Match 92.1%; Score 256; DB 1; Length 87;
 CC Best Local Similarity 92.5%; Pred. No. 5.73e-37;
 CC Matches 37; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
 CC
 Db 48 HGGTF-TSDLSKMEEEAVRLFIEWLNKGGSSGAPPS 86
 QY 1 hsdgtfstdlskmeeeavrlfiewlknngssgppps 40
 CC
 CC ILT 3
 CC GLUC-RANCA STANDARD; PRT; 103 AA.
 AC P15438; P15439; P15440;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE GLUCAGON PRECURSOR (FRAGMENTS).
 OS RANA CATESBEIANA (BULL FROG).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; ANURA;
 CC NEOBATRACHIA; RANOIDEA; RANIDAE; RANINAE; RANA.
 RN [1]
 RP SEQUENCE.
 RC TISSUE=PANCREAS;
 RX MEDLINE; 88257102.
 RA POLLOCK H.G., HAMILTON J.W., ROUSE J.B., EBNER K.E., RAWITCH A.B.:
 RT "Isolation of peptide hormones from the pancreas of the bullfrog
 RT (Rana catesbeiana). Amino acid sequences of pancreatic polypeptide,
 RT oxyntomodulin, and two glucagon-like peptides.";
 RL J. BIOL. CHEM. 263:9746-9751(1988).
 CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -!- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH OTHER SPECIES
 CC SEQUENCES.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 CC PIR; B28091; GCFGB.
 DR PROSITE; PS00250; GLUCAGON; 3.

DR PFAM; PF00123; hormone2; 3.
 DR HSP; P01274; IGCN.
 KW GLUCAGON FAMILY; HORMONE.
 FT PEPTIDE 1 29
 FT PEPTIDE 1 36
 FT PEPTIDE 39 70
 FT NON_CONS 70 71
 FT PEPTIDE 71 103
 SQ SEQUENCE 103 AA; 11719 MW; D43EDFC9 CRC32;
 CC
 CC Query Match 42.4%; Score 118; DB 1; Length 103;
 CC Best Local Similarity 53.1%; Pred. No. 1.29e-07;
 CC Matches 17; Conservative 7; Mismatches 7; Indels 1; Gaps 1;
 CC
 Db 39 HADGTF-TSDSSYLEEKAAREFVDWLKGRP 69
 QY 1 hsdgtfstdlskmeeeavrlfiewlknngp 32
 CC
 CC RESULT 4
 CC ID GLUC-ICTPU STANDARD; PRT; 71 AA.
 AC P04093;
 DT 01-NOV-1986 (REL. 03, CREATED)
 DT 01-NOV-1989 (REL. 10, LAST SEQUENCE UPDATE)
 DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
 DE GLUCAGON PRECURSOR (FRAGMENT).
 OS ICTALURUS PUNCTATUS (CHANNEL CATFISH).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 CC TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; SILURIFORMES; ICTALURIDAE;
 CC ICTALURUS.
 RN [1]
 RP SEQUENCE.
 RC TISSUE=PANCREAS;
 RX MEDLINE; 87156787.
 RA HOSEIN N.M., MAHRENHOLZ A.M., ANDREWS P.C., GURD R.S.:
 RT "Biological activities of catfish glucagon and glucagon-like
 RT peptide.";
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 143:87-92(1987).
 RN [2]
 RP SEQUENCE.
 RC TISSUE=PANCREAS;
 RX MEDLINE; 85157536.
 RA ANDREWS P.C., RONNER P.:
 RT "Isolation and structures of glucagon and glucagon-like peptide from
 RT catfish pancreas.";
 RL J. BIOL. CHEM. 260:3910-3914(1985).
 CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -!- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH AMERICAN
 CC GOOSEFISH SEQUENCES.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 CC PIR; A05166; GCIDC.
 DR PROSITE; PS00260; GLUCAGON; 2.
 DR PFAM; PF00123; hormone2; 2.
 DR HSP; P01274; IGCN.
 KW GLUCAGON FAMILY; HORMONE.
 FT PEPTIDE 1 1
 FT PEPTIDE 1 29
 FT PEPTIDE 38 71
 FT CONFLICT 53 53
 FT NON_TER 71 71
 SQ SEQUENCE 71 AA; 8173 MW; C49ED93A CRC32;
 CC
 CC Query Match 42.1%; Score 117; DB 1; Length 71;
 CC Best Local Similarity 53.1%; Pred. No. 1.99e-07;
 CC Matches 17; Conservative 8; Mismatches 6; Indels 1; Gaps 1;
 CC
 Db 38 HADGTY-TSDVSSYLOEQAAKDFITWLKSGOP 68
 QY 1 hsdgtfstdlskmeeeavrlfiewlknngp 32

-!- X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH AMERICAN GOOSEFISH SEQUENCES.

-!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.

PIR: S06339; GCGXA.

PROSITE: P500260; GLUCAGON; 2.

PFAM: PF00123; hormone2; 2.

HSSP: P01274; IGCN.

GLUCAGON FAMILY; HORMONE.

GLUCAGON.

GLUCAGON-36 (OXYNTOMODULIN).

GLUCAGON-LIKE PEPTIDE.

SEQUENCE 78 AA; 8990 MW; 509ED9D3 CRC32;

Query Match 39.8%; Score 110; DB 1; Length 78;
Best Local Similarity 46.7%; Pred. No. 3,91e-06;
Matches 14; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

Db 45 HDGTY-TSDVSSYLQDAAKKFTVTWKOG 73
YQ 1 hsdgtfidslqkmeeeavrfiewlkng 30

RESULT 7

ID ID GLU2 LOPAM STANDARD; PRI: 122 AA.
AC P04092;

DT 01-NOV-1986 (REL. 03, CREATED)
DT 01-NOV-1986 (REL. 03, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DE GLUCAGON II PRECURSOR.

OS LOPIUS AMERICANUS (AMERICAN GOOSEFISH).
OC BUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
TELOSTEI; EUTELEOSTEI; PARACANTHOPTERYGII; LOPHIIFORMES; LOPHIIDAE;
LOPIUS.

[1]
SEQUENCE FROM N.A.
RX MEDLINE: 83135785.

RA LUND P.K., GOODMAN R.H., MONTMINY M.R., DEE P.C., HABENER J.F.;
RT "Anglerfish islet pre-proglucagon II. Nucleotide and corresponding amino acid sequence of the cDNA.";
RL J. BIOL. CHEM. 258:3280-3284(1983).
RN [2]

RX PROCESSING:
RX MEDLINE: 86286913.
RA NOE B.D., ANDREWS P.C.;
RT "Specific glucagon-related peptides isolated from anglerfish islets are metabolic cleavage products of (pre)proglucagon-II.";
RL PEPTIDES 7:331-339(1986).

-!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOCEN AND LIPIDS, AND RAISES THE BLOOD SUGAR LEVEL.

-!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.

-!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).

EMBL: V00632; G64022; --
DR EMBL: J00933; G213353; --
DR FIR: A05150; GCAP2.
DR PROSITE: P500260; GLUCAGON; 2.
DR PFAM: PF00123; hormone2; 2.
DR HSSP: P01274; IGCN.
GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.

SIGNAL 1 21
FT PEPTIDE 22 49 GRPP (GLICENTINE RELATED POLYPEPTIDE).
FT PEPTIDE 52 80 GLUCAGON II.
FT PEPTIDE 89 119 GLUCAGON-LIKE PEPTIDE II.
SQ SEQUENCE 122 AA; 14171 MW; DF563061 CRC32;


```

QY 1 hsdgtfslskqmeeeavrlfiewlkn 30

RESULT 13
ID GLUC_BOVIN STANDARD; PRT; 180 AA.
AC P01273;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE GLUCAGON PRECURSOR.
GN GCG.

OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; BOVIDAE; BOVINAE; BOS.
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE; 83167563.
CC "Hamster preproglucagon contains the sequence of glucagon and two
CC related peptides."
CC NATURE 302:716-718(1983).
CC [2]
CC REVISIONS TO 12-15.
CC RP
CC BELG G.I.;
CC SUBMITTED (XX-1985) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC THE BLOOD SUGAR LEVEL.
CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J00059; G305354; -
CC PIR: A01539; GCHY.
CC PROSITE; PS00260; GLUCAGON; 4.
CC PFAM; PF00123; hormone2; 3.
CC HSP; P01274; 1GCG.
CC KW GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
CC SIGNAL 1 20
CC FT PEPTIDE 21 50 GRPP (GLICENTINE RELATED POLYPEPTIDE).
CC FT PEPTIDE 53 81 GLUCAGON.
CC FT PEPTIDE 92 128 GLUCAGON-LIKE PEPTIDE 1.
CC FT PEPTIDE 146 178 GLUCAGON-LIKE PEPTIDE 2.
CC FT PEPTIDE 157 178 GLUCAGON-LIKE PEPTIDE 2.
CC SQ SEQUENCE 180 AA; 20954 MW; 126ABD83 CRC32;

Query Match 37.4%; Score 104; DB 1; Length 180;
Best Local Similarity 50.0%; Pred. No. 4.72e-05;
Matches 15; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

Db 98 HAEGTF-TSDVSSYLEGQAAKEFTAWLVKG 126
|:|:| | | | | | | | | | | | | | | | | |
QY 1 hsdgtfslskqmeeeavrlfiewlkn 30

RESULT 15
ID GLUC_MOUSE STANDARD; PRT; 180 AA.
AC P55095;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE GLUCAGON PRECURSOR.
GN GCG.

OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
CC [1]
CC SEQUENCE FROM N.A.
CC TISSUE-PANCREATIC ISLETS;
CC MEDLINE; 95247722.
CC ROTHENBERG M.E.; EILERTSON C.D.; KLEIN K.; ZHOU Y.; LINBERG I.;
CC McDONALD J.K.; MACKIN R.B.; NOE B.D.;
CC "Processing of mouse proglucagon by recombinant prohormone convertase
CC 1 and immunopurified prohormone convertase 2 in vitro."
CC J. BIOL. CHEM. 270:10136-10146(1995).
CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC THE BLOOD SUGAR LEVEL.
CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC -----
CC EMBL: K00107; G163082; -
CC PIR: A01538; GCR0.
CC PROSITE; PS00260; GLUCAGON; 4.
CC PFAM; PF00123; hormone2; 3.
CC HSP; P01274; 1GCG.
CC KW GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
CC SIGNAL 1 20
CC FT PEPTIDE 21 50 GRPP (GLICENTINE RELATED POLYPEPTIDE).
CC FT PEPTIDE 53 81 GLUCAGON.
CC FT PEPTIDE 92 128 GLUCAGON-LIKE PEPTIDE 1.
CC FT PEPTIDE 146 178 GLUCAGON-LIKE PEPTIDE 2.
CC FT PEPTIDE 157 178 GLUCAGON-LIKE PEPTIDE 2.
CC SQ SEQUENCE 180 AA; 20944 MW; 34640341 CRC32;

Query Match 37.4%; Score 104; DB 1; Length 180;
Best Local Similarity 50.0%; Pred. No. 4.72e-05;
Matches 15; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

Db 98 HAEGTF-TSDVSSYLEGQAAKEFTAWLVKG 126
|:|:| | | | | | | | | | | | | | | | | |
QY 1 hsdgtfslskqmeeeavrlfiewlkn 30

RESULT 14
ID GLUC_MESAU STANDARD; PRT; 180 AA.
AC P01273;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE GLUCAGON PRECURSOR.
GN GCG.

OS MESOCRICETUS AURATUS (GOLDEN HAMSTER).

```

```

CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46845; G599881; -
DR MGD; MGI:95674; GCG.
DR PROSITE; PS00260; GLUCAGON; 4.
DR PFAM; PF00123; hormone2; 3.
DR TSSP; P01274; IGCN
DR GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
FT SIGNAL 1 20
FT PEPTIDE 21 50 GRPP (GLICENTINE RELATED POLYPEPTIDE).
FT PEPTIDE 53 81 GLUCAGON.
FT PEPTIDE 92 128 GLUCAGON-LIKE PEPTIDE 1.
FT PEPTIDE 146 178 GLUCAGON-LIKE PEPTIDE 2.
SQ SEQUENCE 180 AA; 20906 MW; 0B21B7BA CRC32;
Query Match 37.4%; Score 104; DB 1; Length 180;
Best Local Similarity 50.0%; Pred. No. 4.72e-05;
Matches 15; Conservative 7; Mismatches 7; Indels 1; Gaps 1;
Db 98 HAEGTF-TSDVSSYLEGQAKKEFTAWLVKG 126
QY 1 hsdgftsdlskqmeeeavrlflewknng 30

```

Search completed: Mon Oct 4 15:29:03 1999
Job time : 8 secs.

M P S R E H
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Oct 4 15:29:20 1999; MasPar time 6.45 Seconds
338.430 Million cell updates/sec

>MOHAM-312-CLAIM83A.PEP
Description: (1-40) from moham312177.ppt
Perfect Score: 278
Sequence: 1 hsdgtfidskqmeeeavrlfiewlknngpssgappps 40

Scoring table: PAM 150
Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sprembl9
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 34.509; Variance 61.367; scale 0.562

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

MULT		Query		Match		Length		DB		ID		Description		Pred. No.	
1	147	52.9	266	13	042143	PROGLUCAGON I.						PROGLUCAGON I.		2.67e-12	
2	142	51.1	219	13	042144	PROGLUCAGON II.						PROGLUCAGON II.		2.38e-11	
3	111	39.9	66	13	091188	GLUCAGON (FRAGMENT)						GLUCAGON (FRAGMENT)		1.03e-05	
4	111	39.9	72	13	091409	PROGLUCAGON (FRAGMENT)						PROGLUCAGON (FRAGMENT)		1.03e-05	
5	111	39.9	72	13	091408	PROGLUCAGON (FRAGMENT)						PROGLUCAGON (FRAGMENT)		1.03e-05	
6	111	39.9	178	13	091971	GLUCAGON I.						GLUCAGON I.		1.03e-05	
7	111	39.9	178	13	091189	GLUCAGON II.						GLUCAGON II.		1.03e-05	
8	103	37.1	206	13	091410	PROGLUCAGON.						PROGLUCAGON.		2.42e-04	
9	99	35.6	149	13	012955	PROGLUCAGON.						PROGLUCAGON.		1.13e-03	
10	99	35.6	204	13	012956	PROGLUCAGON.						PROGLUCAGON.		1.13e-03	
11	95	34.2	2127	14	057294	L. PROTEIN, RNA DEPENDENT						L. PROTEIN, RNA DEPENDENT		5.09e-03	
12	91	32.7	379	2	085863	HYPOTHETICAL 42.3 KD P						HYPOTHETICAL 42.3 KD P		2.24e-02	
13	86	30.9	552	9	038064	REPLICASE.						REPLICASE.		1.36e-01	
14	84	30.2	439	2	067000	APOLIPOPROTEIN N-ACYL						APOLIPOPROTEIN N-ACYL		2.74e-01	
15	82	29.5	1319	2	055359	HYPOTHETICAL 151.9 KD						HYPOTHETICAL 151.9 KD		5.50e-01	
16	81	29.1	338	2	Q59082	UDP-GLUCOSE 4-EPIMERASE						UDP-GLUCOSE 4-EPIMERASE		7.75e-01	
17	80	28.8	2185	3	Q12721	ACETYL COA CARBOXYLASE						ACETYL COA CARBOXYLASE		1.09e+00	
18	79	28.4	502	5	Q22770	T25B9.7 PROTEIN.						T25B9.7 PROTEIN.		1.53e+00	
19	78	28.1	844	5	Q23629	HYPOTHETICAL 95.1 KD P						HYPOTHETICAL 95.1 KD P		2.14e+00	
20	77	27.7	328	2	P96257	HYPOTHETICAL 35.4 KD P						HYPOTHETICAL 35.4 KD P		2.99e+00	

21	77	27.7	412	2	P70868	FILAMENT-A PRECURSOR.						FILAMENT-A PRECURSOR.		2.99e+00	
22	76	27.3	170	9	O80115	L15.						L15.		4.16e+00	
23	76	27.3	300	1	O59151	300AA LONG HYPOTHETICA						300AA LONG HYPOTHETICA		4.16e+00	
24	76	27.3	333	2	Q48538	INTEGRASE/RECOMBINASE.						INTEGRASE/RECOMBINASE.		4.16e+00	
25	76	27.3	610	13	O57319	IAP HOMOLOG.						IAP HOMOLOG.		4.16e+00	
26	76	27.3	691	2	O55726	HYPOTHETICAL 77.8 KD P						HYPOTHETICAL 77.8 KD P		4.16e+00	
27	76	27.3	1141	3	Q74998	NITRITE REDUCTASE (EC						NITRITE REDUCTASE (EC		4.16e+00	
28	76	27.3	1495	11	Q07808	PROTEIN-TYROSINE PHOSP						PROTEIN-TYROSINE PHOSP		4.16e+00	
29	75	27.0	339	2	Q34812	YFMJ PROTEIN.						YFMJ PROTEIN.		5.77e+00	
30	75	27.0	664	2	O59339	ADENYLYL-SULPHATE REDU						ADENYLYL-SULPHATE REDU		5.77e+00	
31	75	27.0	666	11	O35117	PEPTIDYLARGININE DEIMI						PEPTIDYLARGININE DEIMI		5.77e+00	
32	75	27.0	666	11	O88807	PEPTIDYLARGININE DEIMI						PEPTIDYLARGININE DEIMI		5.77e+00	
33	74	26.6	414	5	Q21764	R05H5.4 PROTEIN.						R05H5.4 PROTEIN.		7.98e+00	
34	74	26.6	504	1	O29585	MEDIUM-CHAIN ACYL-COA						MEDIUM-CHAIN ACYL-COA		7.98e+00	
35	73	26.3	159	2	P72565	HYPOTHETICAL 18.1 KD P						HYPOTHETICAL 18.1 KD P		1.10e+01	
36	73	26.3	337	5	O61798	C33E10.8 PROTEIN.						C33E10.8 PROTEIN.		1.10e+01	
37	73	26.3	396	14	O56869	PUTATIVE VIRAL TEGUMEN						PUTATIVE VIRAL TEGUMEN		1.10e+01	
38	73	26.3	418	3	O14418	CARBOXYPEPTIDASE.						CARBOXYPEPTIDASE.		1.10e+01	
39	73	26.3	445	13	P79771	WINGED HELIX PROTEIN C						WINGED HELIX PROTEIN C		1.10e+01	
40	73	26.3	461	5	O01974	CODED FOR BY C. ELEGAN						CODED FOR BY C. ELEGAN		1.10e+01	
41	73	26.3	645	5	O01985	F30F8.8 PROTEIN.						F30F8.8 PROTEIN.		1.10e+01	
42	73	26.3	1015	2	O67146	FORMATE DEHYDROGENASE						FORMATE DEHYDROGENASE		1.10e+01	
43	73	26.3	1352	1	O59837	1352AA LONG HYPOTHETIC						1352AA LONG HYPOTHETIC		1.10e+01	
44	73	26.3	1369	13	O42414	NEUROFASCIN PRECURSOR.						NEUROFASCIN PRECURSOR.		1.10e+01	
45	73	26.3	1756	10	O80975	F26C24.9 PROTEIN.						F26C24.9 PROTEIN.		1.10e+01	

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	266 AA.
ID	O42143			
AC	O42143			
DT	01-JAN-1998 (TREMBLREL. 05, CREATED)			
DT	01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)			
DE	PROGLUCAGON I.			
OS	XENOPUS LAEVIS (AFRICAN CLAWED FROG).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;			
OC	MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 97368292.			
RA	IRWIN D.M., SATKUNARAJAH M., WEN Y., BRUBAKER P.L., PEDERSON R.A.,			
RA	WHEELER M.B.;			
RT	"The xenopus proglucagon gene encodes novel GLP-1-like peptides with			
RT	insulinotropic properties."			
RL	PROC. NATL. ACAD. SCI. U.S.A. 94:7915-7920(1997).			
DR	EMBL; AF004432; G2305016;			
DR	PROSITE; PS00260; GLUCAGON; 5.			
DR	PFAM; PF00123; hormone2; 5.			
SQ	SEQUENCE 266 AA; 30951 MW; E6139A25 CRC32;			

Query Match		52.9%;	Score 147;	DB 13;	Length 266;
Best Local Similarity		57.6%;	Pred. No. 2.67e-12;		
Matches		19;	Conservative	8;	Mismatches
				5;	Indels
					1;
					Gaps
					1;
Db	97	HAEGTF-TSDVTQDLDEKAEKFEIDWLINGGPS 128			
Qy	1	hsdgtfidskqmeeeavrlfiewlknnggps 33			
RESULT		2	PRELIMINARY;	PRT;	219 AA.
ID	O42144				
AC	O42144				
DT	01-JAN-1998 (TREMBLREL. 05, CREATED)				
DT	01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)				
DT	01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)				
DE	PROGLUCAGON II.				
OS	XENOPUS LAEVIS (AFRICAN CLAWED FROG).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;				
OC	MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				

[illegible]

Tue Oct 5 09:37:43 1999

MOHAM-312-CLAIM83A.PEP.ispt

```

Matches 14; Conservative 10; Mismatches 5; Indels 1; Gaps 1;

Db 90 HAGTY-TSDVSYLQDQAKDFVSWLKS 118
  1 hsdgtfitsdlskmeeeavrlfiewlkn 30
  QY 1 hsdgtfitsdlskmeeeavrlfiewlkn 30

RESULT 7
ID Q91189 PRELIMINARY; PRT; 178 AA.
AC Q91189; Q92168;
DT 01-NOV-1996 (TREMBREL. 01, CREATED)
DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE GLUCAGON II.
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIARDNERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;
OC ONCORHYNCHUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-INTESTINE, DISTAL PORTION;
MEDLINE; 95295739.
RT "Trout and chicken proglucagon: alternative splicing generates mRNA
  transcripts encoding glucagon-like peptide 2.";
RL MOL. ENDOCRINOL. 9:267-277(1995).
DR EMBL; U1914; G736363; -.
DR EMBL; U19116; G736372; -.
DR EMBL; U19115; G736372; JOINED.
DR EMBL; U19115; G736371; -.
DR PFAM; PF00123; hormone2; 3.
SQ SEQUENCE 178 AA; 19998 MW; A4299C13 CRC32;

Query Match 39.9%; Score 111; DB 13; Length 178;
Best Local Similarity 46.7%; Pred. No. 1.03e-05;
Matches 14; Conservative 10; Mismatches 5; Indels 1; Gaps 1;

Db 90 HAGTY-TSDVSYLQDQAKDFVSWLKS 118
  1 hsdgtfitsdlskmeeeavrlfiewlkn 30
  QY 1 hsdgtfitsdlskmeeeavrlfiewlkn 30

RESULT 8
ID Q91410 PRELIMINARY; PRT; 206 AA.
AC Q91410;
DT 01-NOV-1996 (TREMBREL. 01, CREATED)
DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE PROGLUCAGON.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95295739.
RA "Trout and chicken proglucagon: alternative splicing generates mRNA
  transcripts encoding glucagon-like peptide 2.";
RL MOL. ENDOCRINOL. 9:267-277(1995).
DR EMBL; S78477; G999387; -.
DR PROSITE; PS00260; GLUCAGON; 3.
DR PFAM; PF00123; hormone2; 3.
SQ SEQUENCE 206 AA; 23875 MW; 8EC91118 CRC32;

Query Match 37.1%; Score 103; DB 13; Length 206;
Best Local Similarity 46.7%; Pred. No. 2.42e-04;
Matches 14; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Db 118 HAEGTY-TSDITSYLEGQAQKEFIAMLVNG 146
  1 hsdgtfitsdlskmeeeavrlfiewlkn 30
  QY 1 hsdgtfitsdlskmeeeavrlfiewlkn 30

```

```

RESULT 9
ID Q12955 PRELIMINARY; PRT; 149 AA.
AC Q12955;
DT 01-JUL-1997 (TREMBREL. 04, CREATED)
DT 01-JUL-1997 (TREMBREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE PROGLUCAGON.
GN LPI.
OS HELODERMA SUSPECTUM (GILA MONSTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;
OC SCLEROGLOSSA; ANGIOMORPHA; HELODERMATIDAE; HELODERMA.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN Y.E., DRUCKER D.J.;
RL J. BIOL. CHEM. 0:0-0(0).
DR EMBL; U77611; G1916063; -.
DR PROSITE; PS00260; GLUCAGON; 1.
DR PFAM; PF00123; hormone2; 2.
SQ SEQUENCE 149 AA; 17224 MW; F763AB51 CRC32;

Query Match 35.6%; Score 99; DB 13; Length 149;
Best Local Similarity 50.0%; Pred. No. 1.13e-03;
Matches 15; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

Db 116 HAGRY-TSDISSYLEGQAQKEFIAMLVNG 144
  1 hsdgtfitsdlskmeeeavrlfiewlkn 30
  QY 1 hsdgtfitsdlskmeeeavrlfiewlkn 30

RESULT 10
ID Q12956 PRELIMINARY; PRT; 204 AA.
AC Q12956;
DT 01-JUL-1997 (TREMBREL. 04, CREATED)
DT 01-JUL-1997 (TREMBREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE PROGLUCAGON.
GN LPII.
OS HELODERMA SUSPECTUM (GILA MONSTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;
OC SCLEROGLOSSA; ANGIOMORPHA; HELODERMATIDAE; HELODERMA.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN Y.E., DRUCKER D.J.;
RL J. BIOL. CHEM. 0:0-0(0).
DR EMBL; U77612; G1916065; -.
DR PROSITE; PS00260; GLUCAGON; 2.
DR PFAM; PF00123; hormone2; 3.
SQ SEQUENCE 204 AA; 23553 MW; EE50250D CRC32;

Query Match 35.6%; Score 99; DB 13; Length 204;
Best Local Similarity 50.0%; Pred. No. 1.13e-03;
Matches 15; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

Db 116 HAGRY-TSDISSYLEGQAQKEFIAMLVNG 144
  1 hsdgtfitsdlskmeeeavrlfiewlkn 30
  QY 1 hsdgtfitsdlskmeeeavrlfiewlkn 30

RESULT 11
ID Q57294 PRELIMINARY; PRT; 2127 AA.
AC Q57294;
DT 01-JUN-1998 (TREMBREL. 06, CREATED)
DT 01-JUN-1998 (TREMBREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBREL. 06, LAST ANNOTATION UPDATE)
DE L PROTEIN, RNA DEPENDENT RNA POLYMERASE.
GN L.
OS RABIES VIRUS.
OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; MONONEGAVIRALES; RHABDOVIRIDAE;
OC LYSSAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RC-HL;
RA MINAMOTO N.;

```


Tue Oct 5 09:37:43 1999

```

RP SEQUENCE FROM N.A.
RC STRAIN-PCC6803;
RX MEDLINE; 97061201.
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
RA SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA TABATA S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. PCC6803. II. Sequence determination of the entire
RT genome and assignment of potential protein-coding regions.";
RL DNA RES. 3:109-136(1996).
DR EMBL; D64003; D1011091; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 1319 AA; 151893 MW; EB2F4ACB CRC32;

Query Match          29.5%; Score 82; DB 2; Length 1319;
Best Local Similarity 34.8%; Pred.No. 5.50e-01;
Matches 9; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Db 594 FLESELVPQLDSEDLVIALDWLKTQG 619
| : | : | : | : | : | : | : | : |
6 fitsdiskqmeeeavrlfiewlknng 31

```

Search completed: Mon Oct 4 15:29:34 1999
 Job time : 14 secs.

```
*****  
M P S R C H  
                (TM)  
*****  
  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
  
mpsrch_pp      protein - protein database search, using Smith-Waterman algorithm  
run on:        Mon Oct 4 15:32:14 1999; MasPar time 13.75 Seconds  
               61.875 Million cell updates/sec  
molecular output not generated.  
  
title:         >MOHAM-312-CLAIM83B.PEP  
description:    (1-40) from moham312177.pep  
perfect score: 281  
sequence:      1 hgegtfidslsgmeeearvlfieklnggpgssgappps 40
```

Post-processing: Minimum Watch 0%				Listing first 45 summaries			
Database:							
a-geneseq35							
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7							
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13							
14:part14 15:part15 16:part16 17:part17 18:part18							
19:part19 20:part20 21:part21 22:part22 23:part23							
24:part24 25:part25 26:part26 27:part27 28:part28							
29:part29 30:part30 31:part31 32:part32 33:part33							
34:part34 35:part35 36:part36 37:part37 38:part38							
39:part39							
Statistics:				Mean 25.274; Variance 106.235; scale 0.238			
pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.							
SUMMARIES							
Result No.	Score	Query Match	Length	ID	Description	Pred. No.	
1	263	93.6	39	39	W61770	Extendin-4, for use in 1.02e-15	
2	263	93.6	39	14	R80546	Heloderma suspectum e 1.02e-15	
3	263	93.6	39	30	W47609	Gila monster extendin- 1.02e-15	
4	263	93.6	87	35	W70288	Heloderma suspectum p 1.02e-15	
5	256	91.1	39	14	R80545	Heloderma horridum ex 4.74e-15	
6	256	91.1	39	39	W61769	Extendin-3, for use in 4.74e-15	
7	256	91.1	39	30	W47608	Gila monster extendin- 4.74e-15	
8	238	84.7	39	39	W61773	Leu(147), Phe(25)-exen 2.40e-13	
9	220	78.3	31	14	R80547	Heloderma suspectum e 1.18e-11	
10	216	76.9	31	14	R80543	Heloderma suspectum e 2.79e-11	
11	208	74.0	30	39	W61771	Extendin-4 (1-30) for 1.56e-10	
12	208	74.0	31	14	R80544	Heloderma suspectum e 1.56e-10	
13	201	71.5	30	29	W39302	H. horridum extendin-4 6.96e-10	
14	196	69.8	30	29	W39309	H. horridum extendin-4 2.02e-09	
15	194	69.0	28	39	W61772	Extendin-4 (1-28) amid 3.09e-09	
16	194	69.0	28	39	W39307	H. horridum extendin-3 3.09e-09	

```
Query Match      93.6%; score 263; DB 39; Length 39;
Best Local Similarity 97.5%; Pred. No. 1.02e-15;
Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

Qy 1 hgegtf-tsldskqmeeeavrlfiewlknpgssgappps 40

RESULT 2

ID R80546 standard; peptide; 39 AA.

AC R80546;

DT 27-FEB-1996 (first entry)

DE Heloderma suspectum extendin-4.

KW Extendin-4; diabetes mellitus; hyperglycaemia; insulinotropic peptide.

OS Heloderma suspectum.

SW US5424286-A.

PD 13-JUN-1995.

PF 24-MAY-1993; 065480.

PR 24-MAY-1993; US-066480.

PA (ENGJ/) ENG J.

PP Eng J;

WP1: 95-262627/34.

Stimulating/inhibiting insulin release with extendin polypeptide(s) - for treating diabetes mellitus and preventing hyperglycaemia.

PS Claim 6; Columns 13-14; 17pp; English.

CC R80546 is Heloderma suspectum extendin-4. It is an insulinotropic peptide, and can therefore be used in the treatment of diabetes mellitus (types I or II), and for the prevention of hyperglycaemia. It normalises hyperglycaemia through glucose-dependent and insulin-(in)dependent mechanisms.

CC and

CC Sequence 39 AA;

Query Match 93.6%; Score 263; DB 14; Length 39;

Best Local Similarity 97.5%; Pred. No. 1.02e-15;

Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 hgegtf-tsldskqmeeeavrlfiewlknpgssgappps 39

Qy 1 hgegtf-tsldskqmeeeavrlfiewlknpgssgappps 40

RESULT 3

ID W47609 standard; peptide; 39 AA.

AC W47609;

DT 03-JUL-1998 (first entry)

DE Gila monster extendin-4.

KW Extendin agonist; gastric motility; gastric emptying; treatment; spasm; postprandial dumping syndrome; postprandial hyperglycaemia; type 1 diabetes; impaired glucose tolerance; toxin ingestion; obesity; Gila monster venom; extendin-4.

KW Heloderma suspectum.

Key Location/Qualifiers

Modified_site 39

/note= "amidated"

PN W09805351-A1.

PD 12-FEB-1998.

PF 08-AUG-1997; U14199.

PR 08-AUG-1996; US-694954.

PA (AMYL-) AMYLIN PHARM INC.

PI Beeley NRA, Gedulin B, Prickett KS, Young AA;

DR WP1: 96-145351/13.

Regulating gastrointestinal motility using extendins or their agonists - for treating spasm, diabetic postprandial hyperglycaemia, impaired glucose tolerance etc., also in diagnostic investigations

PT Claims 20 and 21; Fig 1; 70pp; English.

PS W47549 describes a generic extendin agonist, provided that it does have the formula of either extendin-3 (W47608) or extendin-4 (W47609).

CC Extendin agonists, which reduce gastric motility and delay gastric emptying, can be used to treat spasm (where associated with acute diverticulitis or disorders of the biliary tract or sphincter of Oddi), postprandial dumping syndrome and hyperglycaemia (particularly associated with type 2 diabetes), type 1 diabetes, impaired glucose tolerance, toxin ingestion (an extendin agonist is administered to prevent stomach contents passing into the intestines, then the stomach pumped) and obesity. They can also be administered to subjects undergoing gastrointestinal diagnostic investigation, particularly radiological or by magnetic resonance

CC imaging.

CC Extendins, components of Gila monster venom, have some sequence similarity to glucagon-like peptides (GLP). They are GLP agonists and have been suggested (US5424286) for treatment of diabetes and prevention of hyperglycaemia.

CC Sequence 39 AA;

Query Match 93.6%; Score 263; DB 30; Length 39;

Best Local Similarity 97.5%; Pred. No. 1.02e-15;

Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 hgegtf-tsldskqmeeeavrlfiewlknpgssgappps 39

Qy 1 hgegtf-tsldskqmeeeavrlfiewlknpgssgappps 40

RESULT 4

ID W70288 standard; Protein; 87 AA.

AC W70288;

DT 06-NOV-1998 (first entry)

DE Heloderma suspectum proextendin peptide.

KW Heloderma suspectum proextendin; extendin N-terminal peptide; ENTP; extendin 4 peptide; extendin 3 gene; Heloderma horridum; metabolic disease; drug screening; endocrine tumour; organ failure; cell metabolism; diabetes; reptilian venom peptide.

KW Heloderma suspectum.

OS Heloderma suspectum.

Key Location/Qualifiers

Peptide 1..23

/note= "Signal peptide"

FT Peptide 1..47

/note= "ENTP"

FT Peptide 48..87

/note= "Extendin 4"

FT Cleavage_site 46..47

/note= "Dipeptidyl peptidase cleavage site"

PN W09835033-A1.

PD 13-AUG-1998.

PF 04-FEB-1998; CA0071.

PR 07-FEB-1997; GB-002582.

PR 05-FEB-1997; US-037412.

PA (ONEO-) 1149336 ONTARIO INC.

PI Drucker DJ; Fig 2; 26pp; English.

DR WP1: 98-447230/38.

N-PSDB; V33163.

PT New nucleic acid encoding proextendin - used to diagnose and treat, e.g. endocrine tumours, also to treat poisoning by reptile venom

PS Claim 3; Fig 2; 26pp; English.

CC The Heloderma suspectum proextendin peptide is encoded by its cDNA which was isolated from a H. suspectum salivary gland cDNA library. The proextendin protein comprises of a novel extendin N-terminal peptide (ENTP) linked to the N-terminus of the extendin 4 peptide by a consensus dipeptidyl peptidase cleavage site. The proextendin cDNA can be used to clone or identify related sequences (e.g. the extendin 3 gene of Heloderma horridum, mutant alleles and proextendin gene regulatory defects associated with metabolic disease) and species homologues (e.g. for developing animal models for drug screening).

CC The proextendin peptide can be used to raise antibodies. Anti-proextendin antibodies are claimed to be useful for diagnosing conditions associated with altered levels of proextendin (e.g. endocrine tumours and organ failure) for identifying other regulators of cell metabolism, in drug screens and for treating metabolic diseases (e.g. diabetes) and for neutralising, or detecting, reptilian venom peptides.

CC Sequence 87 AA;

Query Match 93.6%; Score 263; DB 35; Length 87;

Best Local Similarity 97.5%; Pred. No. 1.02e-15;

Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 48 hgegtf-tsldskqmeeeavrlfiewlknpgssgappps 86

Qy 1 hgegtf-tsldskqmeeeavrlfiewlknpgssgappps 40

Tue Oct 5 09:37:44 1999

MOHAM-312-CLAIM83B.PEP.iag

QY 1 hgegtfittsldskqmeeeavrflfiewlknngpssgappps 40

```

RESULT 5
ID R80545 standard; peptide: 39 AA.
AC R80545;
DT 27-FEB-1996 (first entry)
DE Heloderma horridum extendin-3.
KW Extendin-3; diabetes mellitus; hyperglycaemia; insulinotropic peptide.
OS Heloderma horridum.
PN US5424286-A.
PD 13-JUN-1995.
PF 24-MAY-1993; 066480.
PR 24-MAY-1993; US-066480.
PA (ENG/J) ENG J.
PI Eng J; 95-262627/34.
DR WPI; 95-262627/34.
PT Stimulating/inhibiting insulin release with extendin polypeptide(s) -
PT for treating diabetes mellitus and preventing hyperglycaemia.
PS Claim 5; Columns 13-14; 17pp; English.
CC R80545 is Heloderma horridum extendin-3. It is an
CC insulinotropic peptide, and can therefore be used in the treatment of
CC diabetes mellitus (types I or II), and for the prevention of
CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
CC and insulin-(in)dependent mechanisms.
CC Sequence 39 AA;
CC
CC Query Match 91.1%; Score 256; DB 14; Length 39;
CC Best Local Similarity 92.5%; Pred. No. 4.74e-15;
CC Matches 37; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Db 1 hsdgtf-tsdlskqmeeeavrflfiewlknngpssgappps 39
   1:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1 hgegtfittsldskqmeeeavrflfiewlknngpssgappps 40

RESULT 6
ID W61769 standard; peptide: 39 AA.
AC W61769;
DT 29-MAR-1999 (first entry)
DE Extendin-3, for use in treating disorders related to food intake.
KW Extendin; obesity; type II diabetes; eating disorders; cardiac disease;
KW insulin resistance syndrome; elevated plasma glucose level; agonist.
OS Heloderma horridum.
PN W09830231-A1.
PD 16-JUL-1998.
PF 07-JAN-1998; U00449.
PR 14-NOV-1997; US-066029.
PR 08-AUG-1997; US-034905.
PR 14-NOV-1997; US-065404.
PR 14-NOV-1997; US-065442.
PA (AMYL-) AMYLIN PHARM INC.
PI Bealey NRA, Bhavsar S, Prickett KS;
PI WPI; 98-398796/34.
DR Reducing food intake by administering extendins or their
DR analogues - for treatment of e.g. obesity, type II diabetes,
DR eating disorders and insulin resistance
PS Claims 16, 24; Page 8; 214pp; English.
CC The invention relates to a new method for treating disorders that
CC are alleviated by reducing food intake, in particular obesity, type
CC II diabetes, eating disorders, insulin resistance syndrome, elevated
CC plasma glucose levels, or the risk of cardiac disease. The method
CC comprises administering an extendin or an extendin agonist. The treatment
CC reduces appetite and lowers plasma lipid levels. It inhibits food
CC consumption as effectively as amylin or cholecystokinin but has a much
CC longer-lasting action (still effective after 6 hours in a mouse model).
CC The present sequence is that of extendin-3 which is one of the preferred
CC compounds for use in the method.
CC Sequence 39 AA;
CC
CC Query Match 91.1%; Score 256; DB 39; Length 39;
CC Best Local Similarity 92.5%; Pred. No. 4.74e-15;
CC Matches 37; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Db 1 hsdgtf-tsdlskqmeeeavrflfiewlknngpssgappps 39
   1:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1 hgegtfittsldskqmeeeavrflfiewlknngpssgappps 40

```

```

QY 1 hgegtfittsldskqmeeeavrflfiewlknngpssgappps 40

RESULT 7
ID W47608 standard; peptide: 39 AA.
AC W47608;
DT 03-JUL-1998 (first entry)
DE Gila monster extendin-3.
KW Extendin agonist; gastric motility; gastric emptying; treatment;
KW spasm; postprandial dumping syndrome; postprandial hyperglycaemia;
KW type I diabetes; impaired glucose tolerance; toxin ingestion;
KW obesity; Gila monster venom; extendin-3.
OS Heloderma horridum.
PI Bealey NRA, Gedulin B, Prickett KS, Young AA;
PI WPI; 98-145351/13.
DR Regulating gastrointestinal motility using extendins or their
DR agonists - for treating spasm, diabetic postprandial hyperglycaemia,
DR agonists - for treating spasm, diabetic postprandial hyperglycaemia,
DR impaired glucose tolerance etc., also in diagnostic investigations
DR Claims 20 and 21; Fig 1; 70pp; English.
PS W47549 describes a generic extendin agonist, provided that it does
PS have the formula of either extendin-3 (W47608) or extendin-4
PS (W47609).
CC Extendin agonists, which reduce gastric motility and delay gastric
CC emptying, can be used to treat spasm (where associated with acute
CC diverticulitis or disorders of the biliary tract or sphincter of
CC Oddi), postprandial dumping syndrome and hyperglycaemia
CC (particularly associated with type 2 diabetes), type I diabetes,
CC impaired glucose tolerance, toxin ingestion (an extendin agonist is
CC administered to prevent stomach contents passing into the
CC intestines, then the stomach pumped) and obesity. They can also be
CC administered to subjects undergoing gastrointestinal diagnostic
CC investigation, particularly radiological or by magnetic resonance
CC imaging.
CC Extendins, components of Gila monster venom, have some sequence
CC similarity to glucagon-like peptides (GLP). They are GLP agonists
CC and have been suggested (US5424286) for treatment of diabetes and
CC prevention of hyperglycaemia.
CC Sequence 39 AA;
CC
CC Query Match 91.1%; Score 256; DB 30; Length 39;
CC Best Local Similarity 92.5%; Pred. No. 4.74e-15;
CC Matches 37; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Db 1 hsdgtf-tsdlskqmeeeavrflfiewlknngpssgappps 39
   1:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1 hgegtfittsldskqmeeeavrflfiewlknngpssgappps 40

RESULT 8
ID W61773 standard; peptide: 39 AA.
AC W61773;
DT 29-MAR-1999 (first entry)
DE Leu(14), Phe(25)-extendin-4 amide, for reducing food intake.
KW Extendin; obesity; type II diabetes; eating disorders; cardiac disease;
KW insulin resistance syndrome; elevated plasma glucose level; agonist.
OS Heloderma suspectum.
PI Synthetic.
PI Key.
PI Modified_site 39
PI Location/Qualifiers
PI Note- "the C-terminal is in amide form"
PN W09830231-A1.
PD 16-JUL-1998.
PF 07-JAN-1998; U00449.
PR 14-NOV-1997; US-066029.
PR 07-JAN-1997; US-034905.

```

```

PR 08-AUG-1997; US-055404.
PR 14-NOV-1997; US-065442.
PA (AMYL-) AMYLIN PHARM INC.
PI Beeley NRA, Bhavsar S, Prickett KS;
DR WPI; 95-262627/34.
PT Reducing food intake by administering extendins or their
PT analogues - for treatment of e.g. obesity, type II diabetes,
PT eating disorders and insulin resistance.
PS Claims 18, 26; Page 12; 214pp; English.
CC The invention relates to a new method for treating disorders that
CC are alleviated by reducing food intake, in particular obesity, type
CC II diabetes, eating disorders, insulin resistance syndrome, elevated
CC plasma glucose levels, or the risk of cardiac disease. The method
CC comprises administering an extendin or an extendin agonist. The treatment
CC reduces appetite and lowers plasma lipid levels. It inhibits food
CC consumption as effectively as amylin or cholecystokinin but has a much
CC longer-lasting action (still effective after 6 hours in a mouse model).
CC The present sequence is that of an extendin-4 variant which is one of
CC the preferred compounds for use in the method.
SQ Sequence 39 AA;

Query Match 84.7%; Score 238; DB 39; Length 39;
Best Local Similarity 92.5%; Pred. No. 2.40e-13;
Matches 37; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Db 1 hgegtf-tsdlskqleeeavrfliefkngpsgappps 39
QY 1 hgegtfitsdkqmeeeavrfliefkngpsgappps 40

RESULT 9
ID R80547 standard; peptide; 31 AA.
AC R80547;
DE Heloderma suspectum extendin-4 residues 9-39 (Extendin-4(9-39)).
KW Extendin-4; residues 9-39; Extendin-4(9-39);
KW Insulinotropic peptides; inhibitor.
OS Heloderma suspectum.
PN US5424286-A.
PD 13-JUN-1995.
PF 24-MAY-1993; 066480.
PR 24-MAY-1993; US-066480.
PA (ENGJ/) ENG J.
PI Eng J;
PT WPI; 95-262627/34.
PS Stimulating/inhibiting insulin release with extendin polypeptide(s) -
PS for treating diabetes mellitus and preventing hyperglycaemia.
PS Claim 7; Columns 13-14; 17pp; English.
CC R80547 is the Heloderma suspectum extendin-4 residues 9-39. It
CC is an insulinotropic peptide activity inhibitor.
SQ Sequence 31 AA;

Query Match 78.3%; Score 220; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.18e-11;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 dlskqmeeeavrfliefkngpsgappps 31
QY 10 dlskqmeeeavrfliefkngpsgappps 40

RESULT 10
ID R80543 standard; peptide; 31 AA.
AC R80543;
DE Heloderma suspectum extendin-4 residues 1-31 (Extendin-4(1-31)).
KW Extendin-4; residues 1-31; Extendin-4(1-31); diabetes mellitus;
KW hyperglycaemia; insulinotropic peptide.
OS Heloderma suspectum.
PN US5424286-A.
PD 13-JUN-1995.
PF 24-MAY-1993; 066480.
PR 24-MAY-1993; US-066480.
PA (ENGJ/) ENG J.
PI Eng J;
PT WPI; 95-262627/34.
PS Stimulating/inhibiting insulin release with extendin polypeptide(s) -
PS for treating diabetes mellitus and preventing hyperglycaemia.
PS Claim 7; Columns 13-14; 17pp; English.
CC R80543 is the Heloderma suspectum extendin-4 residues 1-31. It is an
CC insulinotropic peptide, and can therefore be used in the treatment of
CC diabetes mellitus (types I or II), and for the prevention of
CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
CC and insulin-(in)dependent mechanisms.
SQ Sequence 31 AA;

Query Match 76.9%; Score 216; DB 14; Length 31;
Best Local Similarity 96.9%; Pred. No. 2.79e-11;
Matches 31; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 hgegtf-tsdlskqmeeeavrfliefknggp 31
QY 1 hgegtfitsdkqmeeeavrfliefknggp 32

RESULT 11
ID W61771 standard; peptide; 30 AA.
AC W61771;
DE 29-MAR-1999 (first entry)
DE Extendin-4 (1-30) for use in treating disorders related to food intake.
KW Extendin; obesity; type II diabetes; eating disorders; cardiac disease;
KW insulin resistance syndrome; elevated plasma glucose level; agonist.
OS Heloderma suspectum.
FH Key Location/Qualifiers
FT Modified_site 30
FT /note= "optionally the C-terminal is in amide form"
PN W09830231-A1.
PD 16-JUL-1998.
PF 07-JAN-1998; U00449.
PR 14-NOV-1997; US-066029.
PR 07-JAN-1997; US-034905.
PR 08-AUG-1997; US-055404.
PR 14-NOV-1997; US-065442.
PA (AMYL-) AMYLIN PHARM INC.
PI Beeley NRA, Bhavsar S, Prickett KS;
DR WPI; 98-398796/34.
PT Reducing food intake by administering extendins or their
PT analogues - for treatment of e.g. obesity, type II diabetes,
PT eating disorders and insulin resistance.
PS Claims 18, 26; Page 11; 214pp; English.
CC The invention relates to a new method for treating disorders that
CC are alleviated by reducing food intake, in particular obesity, type
CC II diabetes, eating disorders, insulin resistance syndrome, elevated
CC plasma glucose levels, or the risk of cardiac disease. The method
CC comprises administering an extendin or an extendin agonist. The treatment
CC reduces appetite and lowers plasma lipid levels. It inhibits food
CC consumption as effectively as amylin or cholecystokinin but has a much
CC longer-lasting action (still effective after 6 hours in a mouse model).
CC The present sequence is that of extendin-4 (1-30) or its amide which is
CC one of the preferred compounds for use in the method.
SQ Sequence 30 AA;

Query Match 74.0%; Score 208; DB 39; Length 30;
Best Local Similarity 96.8%; Pred. No. 1.56e-10;
Matches 30; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 hgegtf-tsdlskqmeeeavrfliefknggp 30
QY 1 hgegtfitsdkqmeeeavrfliefknggp 31

RESULT 12
ID R80544 standard; peptide; 31 AA.
AC R80544;
DE 27-FEB-1996 (first entry)
DE Heloderma suspectum extendin-4 residues 1-31-Tyr31.

```

```

PA (ENGJ/) ENG J.
PI Eng J;
DR WPI; 95-262627/34.
PT Stimulating/inhibiting insulin release with extendin polypeptide(s) -
PT for treating diabetes mellitus and preventing hyperglycaemia.
PS Claim 1; Columns 13-14; 17pp; English.
CC R80543 is the Heloderma suspectum extendin-4 residues 1-31. It is an
CC insulinotropic peptide, and can therefore be used in the treatment of
CC diabetes mellitus (types I or II), and for the prevention of
CC hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
CC and insulin-(in)dependent mechanisms.
SQ Sequence 31 AA;

Query Match 76.9%; Score 216; DB 14; Length 31;
Best Local Similarity 96.9%; Pred. No. 2.79e-11;
Matches 31; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 hgegtf-tsdlskqmeeeavrfliefknggp 31
QY 1 hgegtfitsdkqmeeeavrfliefknggp 32

```


QY 1 hgegtf-tsdlskqmeeeavrflwknng 30

Exendin-4; residues 1-31; Y-31-Exendin-4(1-31); diabetes mellitus;
hyperglycaemia; Tyr31; insulinotropic peptide.
Heloderma suspectum.
US5424286-A.
13-JUN-1995.
24-MAY-1993; 066480.
24-MAY-1993; US-066480.
(ENG/J) ENG J.
Eng J;
WPI; 95-262627/34.
Truncated versions of exendin peptide(s) for treating diabetes -
stimulating/inhibiting insulin release with exendin polypeptide(s) -
for treating diabetes mellitus and preventing hyperglycaemia.
Claim 2; Columns 13-14; 17pp; English.
R050544 is the Heloderma suspectum exendin-4 residues 1-31, where
the native Pro31 has been replaced with a Tyr residue. It is an
insulinotropic peptide, and can therefore be used in the treatment of
diabetes mellitus (types I or II), and for the prevention of
hyperglycaemia. It normalises hyperglycaemia through glucose-dependent
and insulin-(in)dependent mechanisms.
Sequence 31 AA;
Query Match 74.0%; Score 208; DB 14; Length 31;
Best Local Similarity 96.8%; Pred. No. 1.56e-10; Indels 1; Gaps 1;
Matches 30; Conservative 0; Mismatches 0;

Db 1 hgegtf-tsdlskqmeeeavrflwknng 30
QY 1 hgegtf-tsdlskqmeeeavrflwknng 31

RESULT 13
ID W39302 standard; peptide; 30 AA.
AC W39302;
DT 05-JUN-1998 (first entry)
DE H. horridum exendin-4 peptide.
KW Exendin-3; exendin 4; diabetes; insulin; secretion; biosynthesis;
glucagon reduction; hypoglycaemia; glucose; treatment.
OS Heloderma horridum.
FH Key Location/Qualifiers
FT Modified_site 30 /note= "This residue can be any amino acid except Gly"
FT W09746584-A1.
PN 11-DEC-1997.
PD 05-JUN-1997; E02930.
PF 13-SEP-1996; DE-037230.
PR 05-JUN-1996; DE-022502.
PA (BOEF) BOEHRINGER MANNHEIM GMBH.
PI Goeke B, Goeke R, Hoffmann E;
DR WPI; 98-042119/04.
Truncated versions of exendin peptide(s) for treating diabetes -
increase secretion and biosynthesis of insulin, but reduce those of
glucagon, and do not induce hypoglycaemia
Claim 1; Page 4; 150pp; English.
This peptide is a fragment of exendin-4 isolated from Heloderma
horridum. This peptide and its salts, esters and derivatives can be
used to treat diabetes mellitus. They stimulate biosynthesis and
secretion of insulin, but have the opposite effect on glucagon, and
independent of this activity can increase peripheral glucose utilisation.
Exendin-3 and exendin-4 are only active when blood sugar levels are
high, so they will not induce hypoglycaemia. Compared with glucagon-like
peptide 1 (GLP1) and the known exendins, they are more active (effective
at lower doses), more stable to degradation and metabolism and have a
longer lasting effect. Truncated forms of this peptide can be made more
economically than full length versions.
Sequence 30 AA;
Query Match 71.5%; Score 201; DB 29; Length 30;
Best Local Similarity 96.7%; Pred. No. 6.96e-10;
Matches 29; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 hgegtf-tsdlskqmeeeavrflwknng 29
QY 1 hgegtf-tsdlskqmeeeavrflwknng 30

RESULT 14
ID W39309 standard; peptide; 30 AA.
AC W39309;
DT 05-JUN-1998 (first entry)
DE H. horridum exendin-4 peptide derivative #6.
KW Exendin-3; exendin 4; diabetes; insulin; secretion; biosynthesis;
glucagon reduction; hypoglycaemia; glucose; treatment.
OS Heloderma horridum.
FH Key Location/Qualifiers
FT Modified_site 30 /note= "C-terminal amide"

PN W09746584-A1.
PD 11-DEC-1997.
PF 05-JUN-1997; E02930.
PR 13-SEP-1996; DE-037230.
PR 05-JUN-1996; DE-022502.
PA (BOEF) BOEHRINGER MANNHEIM GMBH.
PI Goeke B, Goeke R, Hoffmann E;
DR WPI; 98-042119/04.
Truncated versions of exendin peptide(s) for treating diabetes -
increase secretion and biosynthesis of insulin, but reduce those of
glucagon, and do not induce hypoglycaemia
Claim 2; Page 22; 150pp; English.
Peptides W39303-W39420 are fragments of exendin-3 and exendin-4
isolated from Heloderma horridum which are used in a novel method
for the treatment of diabetes mellitus. These peptides can stimulate
biosynthesis and secretion of insulin, but have the opposite effect on
glucose utilisation. Exendin-3 and exendin-4 are only active when blood
sugar levels are high, so they will not induce hypoglycaemia. Compared
with glucagon-like peptide 1 (GLP1) and the known exendins, they are more
active (effective at lower doses), more stable to degradation and
metabolism and have a longer lasting effect. Truncated forms of this
peptide can be made more economically than full length versions.
Sequence 30 AA;

Query Match 69.8%; Score 196; DB 29; Length 30;
Best Local Similarity 93.3%; Pred. No. 2.02e-09;
Matches 28; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Db 1 hgegtf-tsdlskqmeeeavrflwknng 29
QY 1 hgegtf-tsdlskqmeeeavrflwknng 30

RESULT 15
ID W61772 standard; peptide; 28 AA.
AC W61772;
DT 29-VAR-1999 (first entry)
DE Exendin-4 (1-28) amide for use in treating disorders related to food.
KW Exendin; obesity; type II diabetes; eating disorders; cardiac disease;
insulin resistance syndrome; elevated plasma glucose level; agonist.
OS Heloderma suspectum.
FH Key Location/Qualifiers
FT Modified_site 28 /note= "the C-terminal is in amide form"

PN W09830231-A1.
PD 16-JUL-1998.
PF 07-JAN-1998; U00449.
PR 14-NOV-1997; US-066029.
PR 07-JAN-1997; US-034905.
PR 08-AUG-1997; US-055404.
PR 14-NOV-1997; US-065442.
PA (AMYL-) AMYLIN PHARM INC.
PI Beley NRA, Bhavsar S, Prickett KS;
DR WPI; 98-398796/34.
PT Reducing food intake by administering exendins or their
analogues - for treatment of e.g. obesity, type II diabetes,
eating disorders and insulin resistance
Claims 18, 26; Page 12; 214pp; English.

CC The invention relates to a new method for treating disorders that
CC are alleviated by reducing food intake, in particular obesity, type
CC II diabetes, eating disorders, insulin resistance syndrome, elevated
CC plasma glucose levels, or the risk of cardiac disease. The method
CC comprises administering an exendin or an exendin agonist. The treatment
CC reduces appetite and lowers plasma lipid levels. It inhibits food
CC consumption as effectively as amylin or cholecystokinin but has a much
CC longer-lasting action (still effective after 6 hours in a mouse model).
CC The present sequence is that of exendin-4 (1-28) amide which is one of
CC the preferred compounds for use in the method.

SQ Sequence 28 AA;

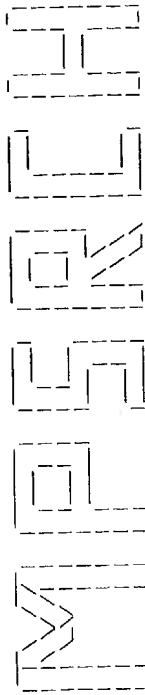
Query Match 59.0%; Score 194; DB 39; Length 28;
Best Local Similarity 96.6%; Pred. No. 3.09e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 hgegtf-tsdlsgmeeeavrllfiewlkn 28

||||| ||||||||| |||||||||

QY 1 hgegtf-tsdlsgmeeeavrllfiewlkn 29

Search completed: Mon Oct 4 15:32:30 1999
Job time : 16 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Oct 4 15:31:48 1999; MasPar time 6.27 Seconds
 255.433 Million cell updates/sec

Abular output not generated.

title: >MOHAM-312-CLAIM83B.PEP
 Description: (1-40) from moham312177.pep
 Perfect Score: 281
 Sequence: 1 hgegtfitsdskmqeeavrlfiewlknngpssgappps 40

Scoring table: PAM 150
Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: pirf60
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 35.094; Variance 64.811; scale 0.541

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	263	93.6	39	1	HWHG4G	1.82e-34
2	256	91.1	39	1	exendin-4 - Gila mons	4.97e-33
3	116	41.3	31	2	HWGH3Z	3.26e-06
4	116	41.3	101	1	glucagon G2 - North A	3.26e-06
5	114	40.6	63	1	glucagon precursor -	7.09e-06
6	111	39.5	30	2	glucagon precursor -	2.55e-05
7	111	39.5	31	2	glucagon-like peptide	2.55e-05
8	110	39.1	30	2	glucagon G1 - North A	3.30e-05
9	110	39.1	30	2	glucagon-like peptide	3.30e-05
10	109	38.8	66	2	glucagon-like peptide	4.82e-05
11	109	38.8	178	2	glucagon - chinook sa	4.82e-05
12	109	38.8	178	2	glucagon II precursor	4.82e-05
13	108	38.4	72	1	glucagon I precursor	7.05e-05
14	107	38.1	122	1	glucagon precursor -	1.04e-04
15	106	37.7	60	1	glucagon precursor -	1.50e-04
16	104	37.0	29	1	glucagon - smaller sp	3.17e-04
17	104	37.0	138	1	glucagon precursor -	3.17e-04
18	104	37.0	180	2	glucagon precursor -	3.17e-04
19	104	37.0	180	1	glucagon precursor -	3.17e-04
20	104	37.0	180	1	glucagon precursor -	3.17e-04
21	104	37.0	180	1	glucagon precursor -	3.17e-04
22	104	37.0	180	1	glucagon precursor -	3.17e-04
23	104	37.0	180	1	glucagon precursor -	3.17e-04

24 104 37.0 180 1 GCRT glucagon precursor - 3.17e-04
 25 103 36.7 151 1 GCCH glucagon precursor - 4.59e-04
 26 103 36.7 206 2 I51301 proglucagon - chicken 4.59e-04
 27 102 36.3 124 1 GCAF glucagon 1 precursor 6.65e-04
 28 99 35.2 29 1 GCCB glucagon - Chinchilla 2.00e-03
 29 97 34.5 29 2 S07211 glucagon - marbled el 4.13e-03
 30 93 33.1 29 1 GCFLE glucagon - European f 1.72e-02
 31 93 33.1 29 2 A61135 glucagon - bigeye tun 1.72e-02
 32 93 33.1 87 1 GCFTS glucagon precursor - 1.72e-02
 33 89 31.7 29 2 A91742 glucagon - Arabian ca 6.97e-02
 34 89 31.7 29 2 A91741 glucagon - rabbit (te 6.97e-02
 35 89 31.7 29 2 C39258 glucagon - common squ 6.97e-02
 36 89 31.7 69 1 GCDG69 glucagon-69 - dog 6.97e-02
 37 88 31.3 29 1 GCEN glucagon - elephantfi 9.84e-02
 38 88 31.3 29 2 S39018 glucagon - bowfin 9.84e-02
 39 88 31.3 2127 1 ZLVNSB genome polyprotein - 9.84e-02
 40 88 31.3 2142 1 ZLVNPFV genome polyprotein - 1.39e-01
 41 87 31.0 29 2 C60840 glucagon 1 - European 1.39e-01
 42 86 30.6 552 2 S46978 replicase - phage PP7 1.95e-01
 43 85 30.2 29 2 A91740 glucagon - turkey (te 2.73e-01
 44 85 30.2 29 2 A91740 glucagon - North Amer 2.73e-01
 45 84 29.9 29 1 GCDK glucagon - duck 3.81e-01

ALIGNMENTS

RESULT 1
 ENTRY HWGH4G #type complete
 TITLE exendin-4 - Gila monster
 ORGANISM #formal_name Heloderma suspectum #common_name Gila monster
 DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 21-Nov-1997
 A42486
 A42486
 Eng, J.; Kleinman, W.A.; Singh, L.; Singh, G.; Raufman, J.P.
 J. Biol. Chem. (1992) 267:7402-7405
 Isolation and characterization of exendin-4, an exendin-3
 analogue, from Heloderma suspectum venom. Further evidence
 for an exendin receptor on dispersed acini from guinea pig
 pancreas.

#cross-references MOID:92218391

#accession A42486

#molecule_type protein

#residues 1-39 #label ENG

COMMENT Exendin-4 does not stimulate amylase secretion by pancreatic acinar
 cells.

CLASSIFICATION #superfamily glucagon

KEYWORDS amidated carboxyl end; duplication; venom

FEATURE 39 #modified_site amidated carboxyl end (Ser) #status experimental

SUMMARY #length 39 #molecular-weight 4188 #checksum 9570

Query Match 93.6%; Score 263; DB 1; Length 39;

Best Local Similarity 97.5%; Pred. No. 1.82e-34;

Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 HGEETF-TSDLSKOMEAEAVRLFIWLNKNGPSSGAPPPS 39

Qy 1 hgegtfitsdskmqeeavrlfiewlknngpssgappps 40

RESULT 2

ENTRY HWGH3Z #type complete

TITLE exendin-3 - Mexican beaded lizard

ORGANISM #formal_name Heloderma horridum #common_name Mexican beaded lizard

DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 21-Nov-1997

A23674

A23674

Eng, J.; Andrews, P.C.; Kleinman, W.A.; Singh, L.; Raufman, J.P.

```

J. Biol. Chem. (1988) 263:9746-9751
Isolation of peptide hormones from the pancreas of the
bullfrog (Rana catesbeiana). Amino acid sequences of
pancreatic polypeptide, oxyntomodulin, and two
glucagon-like peptides.
#cross-references MUID:88257102
#accession B28091
##molecule_type protein
##residues 1-36 ##label PO2
#accession C28091
##molecule_type protein
##residues 37-68 ##label POL
#accession D28091
##molecule_type protein
##residues 69-101 ##label P03
CLASSIFICATION #superfamily glucagon
carbohydrate metabolism; duplication; hormone; pancreas
KEYWORDS
FEATURE
1-36      #product glucagon-36 (oxyntomodulin) #status
experimental #label G36\
1-29      #product glucagon #status predicted #label GCN\
37-67     #product glucagon-like peptide 1 #status experimental
          #label GL1\
69-101    #product glucagon-like peptide 2 #status experimental
          #label GL2
SUMMARY
length 101 #checksum 9108
Query Match      41.3%; Score 116; DB 1; Length 101;
Best Local Similarity 50.0%; Pred. No. 3,26e-06;
Matches 16; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Ddb   37 HADGTF-TSDMSSTLEKAAKEFVDMLIKRP 67
      |:|:| |::|: || |::|:|:|:|:|:|
Qy    1 hgegtfidslqmeeeavrllfewlknngpp 32

RESULT 5
ENTRY      GCIDC      #type fragments
TITLE      glucagon precursor - channel catfish (fragments)
ORGANISM   #formal_name Ictalurus punctatus #common_name channel catfish
DATE       31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
                20-Mar-1998
ACCESSIONS A05166; A05167
REFERENCE   A92514
AUTHORS     Andrews, P.C.; Ronner, P.
JOURNAL     J. Biol. Chem. (1985) 260:3910-3914
TITLES      Isolation and structures of glucagon and glucagon-like
            peptide from catfish pancreas.
#cross-references MUID:85157536
#accession A05166
##molecule_type protein
##residues 1-29 ##label AND1
#accession A05167
##molecule_type protein
##residues 30-63 ##label AND2
CLASSIFICATION #superfamily glucagon
carbohydrate metabolism; duplication; hormone; pancreas
KEYWORDS
FEATURE
1-29      #product glucagon #status experimental #label GCN\
30-63     #product glucagon-like peptide 1 #status experimental
          #label GL1
SUMMARY
length 63 #checksum 9366
Query Match      40.6%; Score 114; DB 1; Length 63;
Best Local Similarity 46.9%; Pred. No. 7,09e-06;
Matches 15; Conservative 10; Mismatches 6; Indels 1; Gaps 1;

Ddb   30 HADGTY-TSDVSSYLDDQAAKDFITWLKSGQP 60
      |:|:| |::|: || |::|:|:|:|:|
Qy    1 hgegtfidslqmeeeavrllfewlknngpp 32

RESULT 6

```

```

ENTRY      S44473      #type complete
TITLE      glucagon-like peptide - North American paddlefish (Polyodon
ORGANISM   spathula)
DATE       18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change
          20-Mar-1998
ACCESSIONS S44473
REFERENCE   S44467
#authors   Nguyen, T.M.; Mommsen, T.P.; Mims, S.M.; Conlon, J.M.
#journal    Biochem. J. (1994) 300:339-345
#title      Characterization of insulins and proglucagon-derived peptides
           from a phylogenetically ancient fish, the paddlefish
           (Polyodon spathula).
#accession  S44473
#molecule_type protein
#residues   1-30 #label NGU
CLASSIFICATION #superfamily glucagon
KEYWORDS      duplication; hormone; pancreas
FEATURE       1-30
           #product glucagon-like peptide #status predicted #label
           MAT
SUMMARY       #length 30 #molecular-weight 3359 #checksum 5186
           39.5%; Score 111; DB 2; Length 30;
           Best Local Similarity 53.3%; Pred. No. 2.25e-05;
           Matches 16; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

Db 1 HADGTY-TSDASSFQEQARDFISWLKKG 29
   |::|: |||: |::|: |::|: |::|: |::|:
   1 hgegtfitsdlskqmeeeavrlfiewlknkng 30

RESULT      7
ENTRY       S44471      #type complete
TITLE      glucagon G1 - North American paddlefish (Polyodon spathula)
ORGANISM    Biochem. J. (1994) 300:339-345
#journal     Characterization of insulins and proglucagon-derived peptides
#title       from a phylogenetically ancient fish, the paddlefish
           (Polyodon spathula).
#accession  S44471
#molecule_type protein
#residues   1-31 #label NGU
#experimental_source pancreas
CLASSIFICATION #superfamily glucagon
KEYWORDS      carbohydrate metabolism; duplication; hormone; pancreas
FEATURE       1-31
           #product glucagon G1 #status predicted #label MAT
SUMMARY       #length 31 #molecular-weight 3751 #checksum 7808
           39.5%; Score 111; DB 2; Length 31;
           Best Local Similarity 53.3%; Pred. No. 2.25e-05;
           Matches 16; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

Db 1 HSQGMF-TNDYSKYLEERAKFEVWLKNG 29
   |::|: |||: |::|: |::|: |::|: |::|:
   1 hgegtfitsdlskqmeeeavrlfiewlknkng 30

RESULT      8
ENTRY       B61125      #type complete
TITLE      glucagon-like peptide - American eel
ORGANISM    #formal_name Anguilla rostrata #common_name American eel
DATE       10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change
          21-Nov-1997
ACCESSIONS B61125
REFERENCE   A61125
#authors   Conlon, J.M.; Andrews, P.C.; Thim, L.; Moon, T.W.
#journal    Mol. Endocrinol. (1995) 9:267-277
#title      Trout and chicken proglucagon: alternative splicing generates

```

```

#journal    Gen. Comp. Endocrinol. (1991) 82:23-32
#title      The primary structure of glucagon-like peptide but not
           insulin has been conserved between the American eel,
           Anguilla rostrata and the European eel, Anguilla anguilla.
#cross-references M01D:91340068
#accession  B61125
#molecule_type protein
#residues   1-30 #label CON
CLASSIFICATION #superfamily glucagon
KEYWORDS      amidated carboxyl end; duplication
FEATURE       1-30
           #product glucagon-like peptide #status experimental
           #label GRP\
           #modified_site amidated carboxyl end (Arg) #status
           predicted
SUMMARY       #length 30 #molecular-weight 3376 #checksum 6092
           39.1%; Score 110; DB 2; Length 30;
           Query Match
           Best Local Similarity 46.7%; Pred. No. 3.30e-05;
           Matches 14; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

Db 1 HAEGTY-TSDVSSYLQDQAAKEFVSWLKTG 29
   |::|: |||: |::|: |::|: |::|: |::|:
   1 hgegtfitsdlskqmeeeavrlfiewlknkng 30

RESULT      9
ENTRY       C61125      #type complete
TITLE      glucagon-like peptide - European eel
ORGANISM    #formal_name Anguilla anguilla #common_name European eel
DATE       10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change
          21-Nov-1997
ACCESSIONS C61125
REFERENCE   A61125
#authors   Conlon, J.M.; Andrews, P.C.; Thim, L.; Moon, T.W.
#journal    Gen. Comp. Endocrinol. (1991) 82:23-32
#title      The primary structure of glucagon-like peptide but not
           insulin has been conserved between the American eel,
           Anguilla rostrata and the European eel, Anguilla anguilla.
#cross-references M01D:91340068
#accession  C61125
#molecule_type protein
#residues   1-30 #label CON
CLASSIFICATION #superfamily glucagon
KEYWORDS      amidated carboxyl end; duplication
FEATURE       1-30
           #product glucagon-like peptide #status experimental
           #label GRP\
           #modified_site amidated carboxyl end (Arg) #status
           experimental
SUMMARY       #length 30 #molecular-weight 3376 #checksum 6092
           39.1%; Score 110; DB 2; Length 30;
           Query Match
           Best Local Similarity 45.7%; Pred. No. 3.30e-05;
           Matches 14; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

Db 1 HAEGTY-TSDVSSYLQDQAAKEFVSWLKTG 29
   |::|: |||: |::|: |::|: |::|: |::|:
   1 hgegtfitsdlskqmeeeavrlfiewlknkng 30

RESULT      10
ENTRY       I51093      #type fragment
TITLE      glucagon - chinook salmon (fragment)
ORGANISM    #formal_name Oncorhynchus tshawytscha #common_name chinook
           salmon
DATE       13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
          21-Nov-1997
ACCESSIONS I51093
REFERENCE   A55895
#authors   Irwin, D.M.; Wong, J.
#journal    Mol. Endocrinol. (1995) 9:267-277
#title      Trout and chicken proglucagon: alternative splicing generates

```

Irwin, D.M.; Wong, J.
Mol. Endocrinol. (1995) 9:267-277

CLASSIFICATION #superfamily glucagon
KEYWORDS carbohydrate metabolism; duplication; hormone; pancreas
FEATURE

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	263	93.6	87	1	EXE4_HELSEU EXENDIN-4 PRECURSOR.	1.95e-38
2	256	91.1	39	1	EXE3_HELHO EXENDIN-3.	7.64e-37
3	116	41.3	103	1	GLUC_RANCA GLUCAGON PRECURSOR (FR	3.34e-07
4	115	40.9	71	1	GLUC_ICTPU GLUCAGON PRECURSOR (FR	5.17e-07
5	110	39.1	30	1	GLUCON-LIKE PEPTIDE	4.29e-06
6	108	38.4	78	1	GLUC_LEPSP GLUCAGON PRECURSOR (FR	9.87e-06
7	108	38.4	121	1	GLUC_CARAU GLUCAGON PRECURSOR.	9.87e-06
8	107	38.1	122	1	GLUC_LOPAM GLUCAGON II PRECURSOR.	1.49e-05
9	106	37.7	68	1	GLUC_ONCKI GLUCAGON II PRECURSOR (FR	2.26e-05
10	104	37.0	129	1	GLUC_SCYCA GLUCAGON.	5.13e-05
11	104	37.0	158	1	GLUC_PIG GLUCAGON PRECURSOR (FR	5.13e-05
12	104	37.0	180	1	GLUC_MOUSE GLUCAGON PRECURSOR.	5.13e-05
13	104	37.0	180	1	GLUC_HUMAN GLUCAGON PRECURSOR.	5.13e-05
14	104	37.0	180	1	GLUC_CAVPO GLUCAGON PRECURSOR.	5.13e-05
15	104	37.0	180	1	GLUC_OCTDE GLUCAGON PRECURSOR.	5.13e-05
16	104	37.0	180	1	GLUC_MESAU GLUCAGON PRECURSOR.	5.13e-05
17	104	37.0	180	1	GLUC_FAT GLUCAGON PRECURSOR.	5.13e-05
18	104	37.0	180	1	GLUC_BOVIN GLUCAGON PRECURSOR.	5.13e-05
19	103	36.7	151	1	GLUC_CHICK GLUCAGON PRECURSOR.	7.72e-05
20	102	36.3	124	1	GLU1_LOPAM GLUCAGON I PRECURSOR.	1.16e-04
21	99	35.2	29	1	GLUC_CHIBR GLUCAGON.	3.87e-04
22	97	34.5	29	1	GLUC_TORMA GLUCAGON II.	8.56e-04
23	94	33.5	33	1	GLU2_ORENI GLUCAGON II.	2.78e-03


```
db      38 HADGTY-TSDVSSYLQEAAKDFITWLKSGQP 68
          |||:||||| :||: ||||:|-|
0y      1 hgeqftisdlskmeeeavrlfiewlknnggp 32
```


RESULT	11	
ID	GLUC PIG	STANDARD;
		PRT; 158 AA.

Query Match 37.0%: Score 104; DB 1; Length 158;

Query Match 37.0%; Score 104; DB 1; Length 180;

Tue Oct 5 09:37:47 1999

MOHAM-312-CLAIM83B.PEP.rsp

Best Local Similarity 53.3%; Pred. No. 5.13e-05; Matches 16; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

Db 98 HAEGTF-TSDVSSYLEGQAAKEFIAMLVKG 126
 QY 1 hgegtfitsdskqmeeeavrlfiewlknq 30

```

RESULT 15
ID GLUC_OCTDE STANDARD; PRT; 180 AA.
AC P22890;
DI 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE GLUCAGON PRECURSOR.
OS OCTODON DEGUS (DEGU).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; HYSTRICOGNATHI; OCTODONTIDAE; OCTODON.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91155952.
QA NISHI M., STEINER D.F.;
  "Cloning of complementary DNAs encoding islet amyloid polypeptide,
  insulin, and glucagon precursors from a New World rodent, the degu,
  Octodon degus.";
  MOL. ENDOCRINOL. 4:1192-1198(1990).
RT -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC -!- THE BLOOD SUGAR LEVEL.
CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC -!- IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M57688; G202468; -
CC PIR; C36118; GCRTDU.
CC PROSITE; PS00260; GLUCAGON; 4.
CC PFAM; PF00123; hormone2; 3.
CC HSP; P01274; 1GCM.
CC GLUCAGON FAMILY; HORMONE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL;
KW AMIDATION.
FT SIGNAL 1 20 GRPP (GLICENTINE RELATED POLYPEPTIDE).
FT PEPTIDE 21 50 GLUCAGON.
FT PEPTIDE 53 81 GLUCAGON-LIKE PEPTIDE 1.
FT PEPTIDE 92 127 GLUCAGON-LIKE PEPTIDE 2.
FT PEPTIDE 146 178 GLUCAGON-LIKE PEPTIDE 2.
FT MOD_RES 127 127 AMIDATION (G-128 PROVIDE AMIDE GROUP).
CC MOD_RES 127 127
CC SEQUENCE 180 AA; 21165 MW; 4A1F5CE3 CRC32;
  
```

Query Match 37.0%; Score 104; DB 1; Length 180;
 Best Local Similarity 53.3%; Pred. No. 5.13e-05; Matches 16; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

Db 98 HAEGTF-TSDVSSYLEGQAAKEFIAMLVKG 126
 QY 1 hgegtfitsdskqmeeeavrlfiewlknq 30

Search completed: Mon Oct 4 15:31:02 1999
 Job time : 6 secs.

M P E R L H
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Oct 4 15:31:19 1999; MasPar time 10.21 Seconds
213.830 Million cell updates/sec
Global output not generated.

Title: >MOHAM-312-CLAIM83B.PEP
Description: (1-40) from moham312177.pep
Perfect Score: 281
Sequence: 1 hgegtfitsdskmeeeavrlfiewlknpgssgappps 40

Scoring table: PAM 150
Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembl9
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 34.596; Variance 60.715; scale 0.570

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match	Length	ID	Description	Pred. No.
1	147	52.3	266	13	PROGLUCAGON I.	1.79e-12
2	142	50.5	219	13	PROGLUCAGON II.	1.64e-11
3	109	38.8	66	13	GLUCAGON (FRAGMENT)	1.86e-05
4	109	38.8	72	13	PROGLUCAGON (FRAGMENT)	1.86e-05
5	109	38.8	72	13	PROGLUCAGON (FRAGMENT)	1.86e-05
6	109	38.8	178	13	PROGLUCAGON I.	1.86e-05
7	109	38.8	178	13	GLUCAGON II.	1.86e-05
8	103	36.7	206	13	PROGLUCAGON.	2.02e-04
9	97	34.5	149	13	PROGLUCAGON.	2.06e-03
10	97	34.5	149	13	PROGLUCAGON.	2.06e-03
11	95	33.8	2127	14	L PROTEIN, RNA DEPENDENT	4.41e-03
12	91	32.4	379	2	HYPOTHETICAL 42.3 KD P	1.97e-02
13	86	30.6	502	5	T25B9.7 PROTEIN.	1.22e-01
14	86	30.6	552	9	REPLICASE	1.22e-01
15	82	29.2	439	2	APOLIPOPROTEIN N-ACYLT	5.02e-01
16	82	29.2	1319	2	HYPOTHETICAL 151.9 KD	5.02e-01
17	80	28.5	2185	3	ACETYL COA CARBOXYLASE	1.00e+00
18	78	27.8	610	13	IAP HOMOLOG.	1.98e+00
19	77	27.4	328	2	HYPOTHETICAL 35.4 KD P	2.78e+00
20	77	27.4	338	2	UDP-GLUCOSE 4-EPIMERAS	2.78e+00

21	77	27.4	412	2	P70868	FILAMENT-A PRECURSOR.	2.78e+00
22	76	27.4	844	5	Q23629	HYPOTHETICAL 95.1 KD P	2.78e+00
23	76	27.0	170	9	O80115	LI5.	3.88e+00
24	76	27.0	300	1	O59151	300AA LONG HYPOTHETICA	3.88e+00
25	76	27.0	333	2	Q48538	INTEGRASE/RECOMBINASE.	3.88e+00
26	76	27.0	691	2	Q55726	HYPOTHETICAL 77.8 KD P	3.88e+00
27	76	27.0	1141	3	O74998	NITRITE REDUCTASE (EC	3.88e+00
28	76	27.0	1495	11	Q07808	PROTEIN-TYROSINE PHOSP	3.88e+00
29	75	26.7	339	2	O34812	YFMJ PROTEIN.	5.40e+00
30	75	26.7	654	2	O59339	ADENYLYL-SULPHATE REDU	5.40e+00
31	75	26.7	666	11	O35117	PEPTIDYLARGININE DEIMI	5.40e+00
32	75	26.7	666	11	O88807	PEPTIDYLARGININE DEIMI	5.40e+00
33	74	26.3	159	2	P72565	HYPOTHETICAL 18.1 KD P	7.49e+00
34	74	26.3	414	5	Q21164	R05H5.4 PROTEIN.	7.49e+00
35	73	26.0	230	2	O67688	HYPOTHETICAL 25.8 KD P	1.04e+01
36	73	26.0	337	5	O61798	C33E10.8 PROTEIN.	1.04e+01
37	73	26.0	396	14	O56859	PUTATIVE VIRAL TEGUMEN	1.04e+01
38	73	26.0	418	3	O14418	CARBOXYPEPTIDASE.	1.04e+01
39	73	26.0	443	5	O62513	ZK550.4 PROTEIN.	1.04e+01
40	73	26.0	445	13	P79771	WINGED HELIX PROTEIN C	1.04e+01
41	73	26.0	504	1	O29585	MEDUFA-CHAIN ACYL-COA	1.04e+01
42	73	26.0	1272	13	Q50924	NEUROFASCIN PRECURSOR.	1.04e+01
43	73	26.0	1352	1	O58837	1352AA LONG HYPOTHETIC	1.04e+01
44	73	26.0	1369	13	O42414	NEUROFASCIN PRECURSOR.	1.04e+01
45	73	26.0	1911	5	O02038	VOLTAGE-GATED CALCIUM	1.04e+01

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	266 AA.
ID	O42143			
AC	O42143;			
DT	01-JAN-1998 (TREMBREL. 05, CREATED)			
DT	01-JAN-1998 (TREMBREL. 05, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)			
DE	PROGLUCAGON I.			
OS	XENOPUS LAEVIS (AFRICAN CLAWED FROG).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;			
OC	MESOBATRACHIA; PIPOIDEA; PIPOIDAE; PIPIDAE; XENOPODINAE; XENOPUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 97368292.			
RA	IRWIN D.M.; SATKUNARAJAH M.; WEN Y.; BRUBAKER P.L.; PEDERSON R.A.;			
RA	WHEELER M.B.;			
RT	"The Xenopus proglucagon gene encodes novel GLP-1-like peptides with			
RT	insulinotropic properties."			
RL	PROC. NATL. ACAD. SCI. U.S.A. 94:7915-7920(1997).			
DR	EMBL; AF004432; G2305016;			
DR	PROSITE; P800260; GLOUCAGON; 5.			
DR	PFAM; PF00123; hormone2; 5.			
SQ	SEQUENCE 266 AA; 30951 MW; E6139A25 CRC32;			

Query Match	52.3%;	Score 147;	DB 13;	Length 266;
Best Local Similarity	60.6%;	Pred. No. 1.79e-12;		
Matches	20;	Conservative	7;	Mismatches 5;
				Indels 1;
				Gaps 1;
Db	97	HAEGTFTSDVTQQLDEKAAKEFIDWLINGSPS 128		
QY	1	hgegtfitsdskmeeeavrlfiewlknpgsps 33		
RESULT	2			
ID	O42144	PRELIMINARY;	PRT;	219 AA.
AC	O42144;			
DT	01-JAN-1998 (TREMBREL. 05, CREATED)			
DT	01-JAN-1998 (TREMBREL. 05, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)			
DE	PROGLUCAGON II.			
OS	XENOPUS LAEVIS (AFRICAN CLAWED FROG).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;			
OC	MESOBATRACHIA; PIPOIDEA; PIPOIDAE; PIPIDAE; XENOPODINAE; XENOPUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			

```

RX MEDLINE; 97368292.
KA IRWIN D.M., SATKUNARAJAH M., WEN Y., BRUBAKER P.L., PEDERSON R.A.,
RA WHEELER M.B.;
RT "The Xenopus prolucagon gene encodes novel GLP-1-like peptides with
RT insulintropic properties."
RC PROC. NATL. ACAD. SCI. U.S.A. 94:7915-7920(1997).
DR EMBL; AF004433; G2305018; -.
DR PROSITE; PS00260; GLUCAGON; 3.
DR PFAM; PF00123; hormone2; 4.
SQ SEQUENCE 219 AA; 25271 MW; 45042A88 CRC32;

Query Match 50.5%; Score 142; DB 13; Length 219;
Best Local Similarity 54.5%; Pred. No. 1.86e-11;
Matches 18; Conservative 9; Mismatches 5; Indels 1; Gaps 1;

Db 97 HADGTY-TSDVSTYLOQAADKDFVSLKSG 128
1 hgegtfitsdskmeeeavrlfiewlknngps 33

RESULT 3
ID Q91188 PRELIMINARY; PRT; 66 AA.
AC Q91188;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE GLUCAGON (FRAGMENT).
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIRONERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;
OC ONCORHYNCHUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95295739.
RA IRWIN D.M., WONG J.;
RT "Trout and chicken prolucagon: alternative splicing generates mRNA
RT transcripts encoding glucagon-like peptide 2."
RL MOL. ENDOCRINOL. 9:267-277(1995).
DR EMBL; U19913; G736361; -.
DR PFAM; PF00123; hormone2; 2.
FT NON-TER 1
SQ SEQUENCE 66 AA; 7680 MW; 62C575E2 CRC32;

Query Match 38.8%; Score 109; DB 13; Length 66;
Best Local Similarity 43.3%; Pred. No. 1.86e-05;
Matches 13; Conservative 11; Mismatches 5; Indels 1; Gaps 1;

Db 33 HADGTY-TSDVSTYLOQAADKDFVSLKSG 61
1 hgegtfitsdskmeeeavrlfiewlknng 30

RESULT 4
ID Q91409 PRELIMINARY; PRT; 72 AA.
AC Q91409;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE PROGLUCAGON (FRAGMENT).
OS ONCORHYNCHUS TSCHAWYTSCHA (CHINOOK SALMON) (KING SALMON).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;
OC ONCORHYNCHUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95295739.
RA IRWIN D.M., WONG J.;
RT "Trout and chicken prolucagon: alternative splicing generates mRNA
RT transcripts encoding glucagon-like peptide 2."
RL MOL. ENDOCRINOL. 9:267-277(1995).
DR EMBL; U19920; G736367; -.

```

```

DR PFAM; PF00123; hormone2; 2.
FT NON-TER 1
SQ SEQUENCE 72 AA; 8293 MW; 0F7AF3EC CRC32;

Query Match 38.8%; Score 109; DB 13; Length 72;
Best Local Similarity 43.3%; Pred. No. 1.86e-05;
Matches 13; Conservative 11; Mismatches 5; Indels 1; Gaps 1;

Db 39 HADGTY-TSDVSTYLOQAADKDFVSLKSG 67
1 hgegtfitsdskmeeeavrlfiewlknng 30

RESULT 5
ID Q91408 PRELIMINARY; PRT; 72 AA.
AC Q91408;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE PROGLUCAGON (FRAGMENT).
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIRONERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;
OC ONCORHYNCHUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95295739.
RA IRWIN D.M., WONG J.;
RT "Trout and chicken prolucagon: alternative splicing generates mRNA
RT transcripts encoding glucagon-like peptide 2."
RL MOL. ENDOCRINOL. 9:267-277(1995).
DR EMBL; S78473; G999383; -.
DR PFAM; PF00123; hormone2; 2.
FT NON-TER 1
SQ SEQUENCE 72 AA; 8293 MW; 0F7AF3EC CRC32;

Query Match 38.8%; Score 109; DB 13; Length 72;
Best Local Similarity 43.3%; Pred. No. 1.86e-05;
Matches 13; Conservative 11; Mismatches 5; Indels 1; Gaps 1;

Db 39 HADGTY-TSDVSTYLOQAADKDFVSLKSG 67
1 hgegtfitsdskmeeeavrlfiewlknng 30

RESULT 6
ID Q91971 PRELIMINARY; PRT; 178 AA.
AC Q91971;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE GLUCAGON 1.
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIRONERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES; SALMONIDAE;
OC ONCORHYNCHUS.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-INTESTINE, DISTAL PORTION;
RX MEDLINE; 95295739.
RA IRWIN D.M., WONG J.;
RT "Trout and chicken prolucagon: alternative splicing generates mRNA
RT transcripts encoding glucagon-like peptide 2."
RL MOL. ENDOCRINOL. 9:267-277(1995).
DR EMBL; U19917; G736365; -.
DR EMBL; S78475; G999385; -.
DR PROSITE; PS00260; GLUCAGON; 3.
DR PFAM; PF00123; hormone2; 3.
SQ SEQUENCE 178 AA; 20034 MW; 2056F963 CRC32;

Query Match 38.8%; Score 109; DB 13; Length 178;
Best Local Similarity 43.3%; Pred. No. 1.86e-05;

```


RC STRAIN-VF5;
RA DECKERT G., WARREN P.V., GASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AOJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AE000709; G2983374; -;
KW TRANSFERASE; ACYLTRANSFERASE; LIPOPROTEIN.
SQ SEQUENCE 439 AA; 50757 MW; 7963CD20 CRC32;

Query Match 29.2%; Score 82; DB 2; Length 439;
Best Local Similarity 35.1%; Pred. No. 5.02e-01;
Matches 13; Conservative 10; Mismatches 13; Indels 1; Gaps 1;

Db 384 SEGTECHMKLAVRALENEKEFL-WYNTGPGGIISP 419
QY 2 gegtftslskmqeeavrlfiewlknngpsgapp 38

Search completed: Mon Oct 4 15:31:31 1999
Job time : 12 secs.